12

STUDY OF THE PHYSICS OF INSULATING FILMS AS RELATED TO THE RELIABILITY OF METAL-OXIDE SEMICONDUCTOR DEVICES

C. Aliotta	C. Falcony	W. Reuter
J. Blum	F.J. Feigl	D.J. Robbins
S.D. Brorson	B. Ginsberg	C.M. Serrano
J. Calise	F.F. Fang	P. Solomon
I.F. Chang	C.R. Guarnieri	T. Theis
J.F. DeGelormo	A. Hartstein	E. Tierney
K.M. DeMeyer	E. Irene	J.C. Tsang
D.J. DiMaria	J.R. Kirtley	Z.A. Weinberg
D.W. Dong	S.K. Lai	S. Yokoyama
D.B. Dove	A. Lamberti	D.R. Young
L.M. Ephrath	F.L. Pesavento	

DTIC
SELECTED
OCT 18 1983
D

Approved for public release; distribution unlimited

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the United States Government.

83 10 17 068

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A133 674	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STUDY OF THE PHYSICS OF INSULATING FILMS AS RELATED TO THE RELIABILITY OF METAL-OXIDE SEMICONDUCTOR DEVICES		5. TYPE OF REPORT & PERIOD COVERED FINAL TECHNICAL REPORT 1 Dec 80 - 30 Nov 82
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) D.R. Young et al. (See complete list included in report)		8. CONTRACT OR GRANT NUMBER(s) MDA903-81-C-0100 ARPA No. 4102
9. PERFORMING ORGANIZATION NAME AND ADDRESS IBM Thomas J. Watson Research Center P.O. Box 218 Yorktown Heights, NY 10598		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE August 1983
		13. NUMBER OF PAGES 244
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Deputy for Electronic Technology Rome Air Development Center (RADC/ESR) Hanscom MA 01731		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: John C. Garth (RADC/ESR)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Light emission, hot electron distribution, Si-rich SiO_2, electron trapping, ellipsometry, Si oxidation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The papers enclosed with this report include: a new method for studying hot electron energy distributions in SiO_2, plasma enhanced chemical vapor deposition of Si-rich SiO_2, the use of Si-rich SiO_2 to greatly reduce electron trapping effects, the use of Si-rich SiO_2 to increase the yield of thin film capacitors, ellipsometry measurements of polycrystalline silicon films and the use of a delay time technique to measure the diffusion of the oxidant in SiO_2 films.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

BLANK PAGE

II. SUMMARY OF WORK PERFORMED

In the discussion that follows the numbers in parenthesis refer to the listing of papers contained in Section III.

The continuing rapid evolution of high density integrated circuits has required a careful examination and control of all the vital materials involved. The thin insulating layer (SiO_2) used for MOSFET's is a crucial material that plays a dominant role in modern high density integrated circuit technology. Projections for future extensions of the technology indicate the need to have the capability of making thin (10 nm) layers of this material with improved yield and reliability. In addition non volatile memory devices are utilizing complex structures imbedded in this materal. The accomplishments of this project have been centered around these requirments.

Our continued work with the Si rich SiO_2 has enhanced our conviction that this material offers the greatest possibilities for application to charge injectors in the non volatile devices of the future. We have made the Si rich material in three different IBM locations, using atmospheric chemical vapor deposition processes, low pressure chemical vapor deposition processes and rf plasma assisted deposition processes. The basic characteristics of the material have been reproduced in all these cases, although minor differences exist, it is evident that the process chosen for final use can depend on manufacturing considerations. This clearly demonstrates the reproducibility of the characteristics that we have previously described of this material. The IBM Raleigh Laboratory has made 36k bit fully functional chips. This clearly demonstrates the application of the technology to large scale memory applications and shows that suitable cell designs can be realized that are compatible with the peripheral circuits circuits located on the same chip. One of the most important features of this development is the use of tooling that already exists in the semiconductor manufacturing plants.

The Double Injector Structures that we have proposed enable us to write and erase the floating gates electronically. Some previous designs required other techniques for erasure. They also avoid the high fields required for the usual tunnel oxide structures. We believe this avoids problems associated with yield and reliability that occur with high field devices. It has been gratifying to observe that the enhanced injection of the Si rich charge injectors only plays a role when the fields required for writing and erasing are used and that when the lower fields used for normal device operation are applied the effective barrier height is the same as for the Si-SiO$_2$ interface. This results in the same charge loss characteristic as for the usual floating gate structure. As a result, the charge retention characteristics are excellent and we do not have to make the usual compromises associated with MNOS Devices.

The number of useful write-erase cycles for floating gate structures is limited by electron trapping in the oxide dielectric. This limits the number of these cycles to be about 10^5 to 10^6. Since the electron trapping behavior of the usual oxides is controlled by the content of the water related species, this can be improved somewhat by improvments in this respect, never the less large improvments had not seemed likely. It also seemed unlikely that a better insulator than SiO$_2$ could be found since the electron trapping rate in SiO$_2$ is already very low in comparison with other known materials. As result it was completely unexpected by us that a substantial improvment could be made. Never the less, it turned out that, as a result of a suggestion by one of our people (David Dong), a substantial improvment has been made in this respect. With this improved material it has been possible to increase the number of useful cycles from 10^6 to 10^{10}. This is accomplished by using Si rich SiO$_2$ material for the intervening oxide as well as for the injectors. In this case the Si content is lower than used in the injectors. The drastic reduction in the trapping rate is confirmed by trapping studies as well as by the device cyclability studies. Although the mechanism for this large improvment is still under investigation, two possibilities are being considered:

1. Since the current in this material is largely carried by electrons with a tunneling process, these electrons are not as effectively trapped as electrons flowing through the conduction band of the SiO_2.

2. The presence of the excess silicon particles in the vicinity of the normal traps provides sites that the trapped electrons can tunnel into and as a result they do not build up charges in the insulator.

We are excited about this recent observation and we expect this will open up new opportunities for use of these devices that would not be feasible otherwise. It would be expected that the use of too much excess Si would degrade the charge retention characteristics and this is the case; however it has been observed that a useful range exists before this problem becomes excessive. This work briefly described above is discussed in the papers by D.J. DiMaria, K.M. DeMeyer, C.M. Serrano, D.W. Dong (14), by D.J. DiMaria, D.W. Dong, C. Falcony, S.D. Brorson (22), and by D.J. DiMaria, D.W. Dong, F.L. Pesavento (30).

We have been able to use our charge injectors for other practical applications in addition to the application to non volatile memory devices. This arises due to the possibility of injecting large current densities (10-100 amp/cm^2) into the SiO_2. One application of this type is the use of these injectors in conjunction with an electroluminescent material to make an efficient device for emitting light that operates with the application of a dc voltage. It is realized that many future applications of this type will develop in the future. The application of the Si rich technology to these devices is described in the paper by Robbins, D.J. DiMaria, C. Falcony, D.W. Dong (26).

Our experience with the reproducibility of the injecting characteristics of our charge injectors has suggested another application involving the storage capacitor of a dynamic memory device. The use of a soft contact between the electrode and the SiO_2 greatly reduces the incidence of low field breakdowns and thus significantly increases the yield of useful

devices. Another interesting observation concerning this application is the elimination of destructive breakdown as a result of the passage of large currents through the device which frequently occurs with the absence of the soft contact. This contact works best if it can be used at the negative electrode. Studies of this application are reported in the paper by S.K. Lai, D.J. DiMaria, F.F. Fang (29).

In addition, the ability to induce these large current flows through the SiO_2 has made possible scientific investigations that were previously impossible. One of the investigations of this type is concerned with the energy spectrum of hot electrons in SiO_2. This spectrum is of considerable practical interest due to the need for a better understanding of electrical break-down mechanisms is SiO_2. One of the techniques that we have used to study the energy distribution is to measure the spectrum of the surface plasmons generated when electrons fall from the conduction band of SiO_2 into the Fermi level of the gate electrode. This work would be impossible with out the ability to induce the large current densities in the SiO_2. Our results to date indicate that a considerable electron heating occurs due to the electric field applied which is contrary to a number of theories that have been suggested for SiO_2. This new work has been discussed in the published papers by Theis, Kirtley, DiMaria, Dong (24), and by Kirtley, Theis, Tsang, DiMaria (25).

The combination of the ultra sensitive IR total reflectance technique with our electron trapping measurments has enabled us, Hartstein and Young, (1) to identify some traps that had been previously observed. These traps are related to water related species in the oxide. It had been observed that the cross section of the dominant trap in a sample without a post metallization annealing treatment was 1×10^{-17} which decreased to 2×10^{-18} after a 400°C post metallization treatment. The IR studies identified these traps as Si-OH and loosely bound H_2O respectively. These results are in keeping with our long range goals to gain a microscopic understanding of the traps we have been studying.

The developments in increasing the density of Si integrated circuits are placing increasing importance on a continual decrease in the thickness of the insulator (SiO_2) used. One of the limitations in the minimum useful thickness is the increase in the relative importance of the distribution of devices with a premature electrical breakdown. As a result of considerable work done at IBM we have concluded that electron traps play a vital role in improving breakdown characteristics by electrically screening out the effects of defects at the conductor-insulator interfaces. This suggests that the problem associated with a thin insulator may be due to the small number of traps contained in the bulk of this material since the total number of traps decreases as the thickness decreases. Considerations of this type are discussed in the paper by Lai (2).

A major concern has been the damage to the insulator resulting from the radiation applied during the construction of the device. One process of interest in this respect is Reactive Ion Etching. The papers by Ephrath and DiMaria (3) and by Ephrath, DiMaria and Pesavento (11) are concerned with this problem and discuss means for minimizing this effect. The effect of the processing environment on the electron trapping characteristics of SiO_2 has been the basis of considerable work in the past and continues to be important. We have observed that the use of very careful processing conditions can decrease the electron trapping characteristics of SiO_2 by almost two orders of magnitude. The precautions required to achieve this result are concerned with the oxidation procedure and require the reduction of the water related species that are inadvertantly added to the oxide during the oxidation procedure. This work is described in the paper by Lai, Young, Calise and Feigl (8).

Electron trapping in SiO_2 is dominated by the water related species that are inadvertantly incorporated into the oxide. The paper previously referred to shows that large reductions in the trapping rate can be made if precautions are taken to reduce the concentration of these species. To learn more about the role of water we have deliberately exposed the SiO_2 to a low temperature (100-300°C) water vapor ambient to learn more about the effect of water

diffusion on the trapping characteristics. The results of this work are described in the paper by Feigl, Young, DiMaria, Lai and Calise (7). This paper also describes the role of water species on the generation of donor states at the Si-SiO$_2$ interface.

As the thickness of the oxide layer decreases the relative importance of interface effects increases. As a result, it becomes increasingly important to gain an increased understanding of interface effects. One type of effect is concerned with the buildup of slow states (donor states) at the SiO$_2$ interface as a result of the passage of a current through the oxide for a prolonged period of time. This effect had been previously observed and discussed by IBM workers. New work concerning these slow states is discussed in the papers by Feigl, Young, DiMaria, Lai (4), by Lai, Young (5), by Feigl, Young, DiMaria, Lai, Calise (7), and by Lai, Young (9). In addition to the generation of slow states, fast states are also generated. These fast states are described in the papers by Lai and Young (5), by Lai (6), and by Lai and Young (9).

In conjunction with some of the IBM manufacturing activities it was observed that there was an anomalous charge located at an interface between SiO$_2$ and Si$_3$N$_4$ resulting from an impurity remaining from a wafer cleaning step. Further investigation revealed that this impurity was Al. It was postulated that this impurity was originally on the surface of the silicon and that as a result of oxidation this impurity was transported to the outside surface of the oxide. In other words, the oxident was transported through the Al (probably Al$_2$O$_3$) and the oxide was grown underneath. When the subsequent Si$_3$N$_4$ was applied by a CVD process the Al remained at the interface. This effect was simulated and described in the paper by DiMaria, Reuter, Young, Pesavento, and Calise (16). In this case an approximately monomolecular layer of Al was evaporated on one half of a silicon wafer. An oxide layer was thermally grown and then a SiO$_2$ layer was deposited by a CVD process. The portion of the wafer with the Al had an electron trapping rate that was 2-3 orders of magnitude greater than the portion of the wafer without th.. In addition it was found, using photo I-V measurments, that the

trapped charge was located at the interface between the thermal SiO_2 and the CVD SiO_2. In addition SIMS measurments located the Al at this interface. This work clearly demonstrates the sensitivity of structures with interfaces of this type to impurities residing on the Silicon wafer.

The statement of work proposed to be performed under this contract is as follows:

1. Continue our basic studies of electron and hole trapping in SiO_2, investigation of radiation damage effects encountered during device construction, and measurments of the effect of various annealing procedures with emphasis on thin insulators ($\sim$ 100 Å).

2. Extend our measurement capability to enable us to evaluate thin-high dielectric constant gate insulators.

3. Develop technologies for making useful thin-high dielectric constant gate insulators.

4. Measure barrier height for oxynitride films and evaluate hot electron injection for practical devices using these films.

5. Study radiation damage effects for thin-high dielectric constant gate insulators and determine the effect of various annealing treatments to eliminate this damage.

6. Investigate means to reduce the voltage required for charge injection using charge injector structures.

7. Evaluate other materials for charge injectors (at the present time we are using Si rich SiO_2) and compare CVD techniques with RF plasma techmiques as means for making these materials.

8.   Continue basic studies of charge injection mechanisms responsible for operation of charge injectors.

9.   Study the effect of a large number ($\sim 10^5$) of write-erase cycles on the EAROM structures using charge injectors and develop means to increase the number of useful write-erase cycles.

CONCERNING THE ORIGINAL PROPOSAL

1.   We have continued our study of electron trapping in SiO_2, we have identified the chemical nature of the water related species, we have studied the effect of radiation coming from reactive ion etching and developed means to minimize the damage created in the SiO_2 and an annealing procedure has been developed that results in an oxide with an ultra low electron trapping rate.

2.   We have worked on the problems associated with measurments of the properties of thin insulators and realize the increasing importance of the $Si-SiO_2$ interface on the resultant characteristics. Bulk effects are not completely dominant as in the case of thick insulators.

3.   We have applied the Si rich technology to thin insulator structures and have shown that this technology can improve the yield of devices using thin insulators.

4.   The work on oxynitride has not progressed as expected.

5.   The radiation damage work has been concentrated on the effect of reactive ion etching which has been studied extensively with conclusions that the damage created is not excessive if reasonable precautions are taken.

6.   The voltage for charge injection in charge injector structures has been reduced by determining the optimum Si rich composition and by decreasing the oxide thickness. As a

result it has been possible to make fully functional circuits using this technology with integrated circuits for addressing and readout as well as for the storage function.

7.   Cermets using metal particles have been investigated as a possible alternate technology to Si rich SiO_2. These materials also work but do not offer sufficient advantages to warrent further consideration. The use of Si rich material offers the major advantage of using existing manufacturing tooling and therefore a departure from this technology would only be made if a significant improvment resulted which did not occur.

8.   Papers have been published describing the basic operation of the Si rich charge injectors as particle-particle tunneling with field enhanced injection into the SiO_2.

9.   Materials have been developed for the intervening oxide (also Si rich SiO_2) that enable us to increase the number of read-write erase cycles from 10^5 to 10^{10}.

We have also included work that was not originally proposed since it resulted from unexpected developments. This work includes the evaluation of a new Electroluminescent device, light emission from Tunnel and Electron Injector Structures, and work on diffusion of the oxidant in SiO_2 using a delay time technique.

MAJOR ACCOMPLISHMENTS

1.   Water related traps identified in SiO_2.

2.   Evaluated damage to SiO_2 resulting from radiation generated by reactive ion etching process.

3.   Studied two step process associated with hole trapping and electron capture which generates surface states at the Si-SiO_2 interface.

4.   Explored the effect of water diffusion on electron trapping in SiO_2 and interface state generation at the $Si-SiO_2$ interface.

5.   Model developed for double injector structure using Si rich charge injectors.

6.   Extended cyclability of electrically alterable Si-rich charge injector structures from 10^5 to 10^{10}.

7.   Measured light emission from tunnel junctions and from charge injectors as a new technique for studying hot electron energy distributions in SiO_2.

8.   Made a new electroluminescent device using Si rich charge injectors.

III. PAPERS PUBLISHED UNDER THIS CONTRACT

1.   "Identification of Electron Traps in Thermal Silicon Dioxide Films", A. Hartstein and D.R. Young, Appl. Phys. Lett. 38, 631 (1981).

2.   "Very Thin Insulators", S.K Lai, in Semiconductor Silicon 1981, edited by H.R. Huff, R.J. Kriegler, and Y. Takeishi (The Electrochemical Society Inc., N.J., 1981), PP. 416-423.

3.   "Review of RIE Induced Radiation Damage in Silicon Dioxide", L.M. Ephrath, D.J. DiMaria, Solid State Technology, P. 188, April 1981.

4.   "Oxide and Interface Charge Generation by Electron Injection on MOS Devices", F.J. Feigl, D.R. Young, D.J. DiMaria and S.K. Lai, Insulating Films on Semiconductors, edited by M. Schulz and G. Pensl (Springer-Verlag Berlin Heidelberg, 1981) pp. 104-110.

5.   "Interface Effects in Avalanche Injection of Electrons into Silicon Dioxide". S.K. Lai and D.R. Young, edited by M. Schulz and G. Pensl (Springer-Verlag Berlin Heidelberg, 1981) pp 118-121.

6.   "Two Carrier Nature of Interface State Generation in Hole Trapping and Radiation Damage", S.K. Lai, Appl. Phys. Lett. 39, 58 (1981).

7.   "The Effects of Water on Oxide and Interface Trapped Charge Generation in Thermal Silicon Dioxide Films", F.J. Feigl, D.R. Young, D.J. DiMaria, S.K. Lai, and J.A. Calise. J. Appl. Phys. 52, 5665 (1981).

8.   "Reduction of electron trapping in silicon dioxide by high- temperature nitrogen anneal". S.K. Lai, D.R. Young, J.A. Calise, and F.J. Feigl, J. Appl. Phys., 52, 5691 (1981).

9.   "Effects of avalanche injection of electrons into silicon dioxide- generation of fast and slow interface states", S.K. Lai and D.R. Young, J. Appl. Phys. 52, 6231 (1981).

10.   "Electron Microscope Investigation of the Effect of Phosphorus Doping on the Plasma Etching of Polycrystalline Silicon", E.A. Irene, E. Tierney, J. Blum, C. Aliotta, A. Lamberti, and B. Ginsbe , J. Electrochem. Soc. 128, 2415 (1981).

11.   "Parameter Dependence of RIE Induced Radiation Damage in Silicon Dioxide", L.M. Ephrath, D.J. DiMaria, and F. Pesavento, J. Electrochem. Soc. 128, 2415 (1981).

12.   "Effect of Forming Gas Anneal on Al-SiO$_2$ Internal Photoemission Characteristics", P.M. Solomon and D.J. DiMaria. J. Appl Phys. 52, 5867 (1981).

13.   "Dual-Electron-Injector-Structure Electrically-Alterable Read- Only-Memory Modeling Studies", D.J. DiMaria, K.M. DeMeyer and D.W. Dong, IEEE Trans. on Electron Devices ED-28, 1047 (1981).

14. "Electrically-Alterable Read-Only-Memory Using Si-Rich SiO_2 Injectors and a Floating Polycrystalline Silicon Storage Layer", D.J. DiMaria, K.M. DeMeyer, C.M. Serrano, and D.W. Dong, J. Appl. Phys. 32, 4825 (1981).

15. "Charge Injection into Wide Energy Band-Gap Insulators", D.J. DiMaria, in Insulating Films on Semiconductors, edited by M. Schulz and G. Pensl, (Springer-Verlag Berlin Heidelberg, 1981) pp. 88-103.

16. "Detection of Inpurities on Silicon Surfaces", D.J. DiMaria, W. Reuter, D.R. Young, F.L. Pesavento, and J.A. Calise, J. Appl. Phys. 52, 1366 (1981).

17. "Capture and Release of Electrons on Na^+ Related Trapping Sites in the SiO_2 Layer of Metal-Oxide-Semiconductor Structures at Temperatures Between 77 K and 296 K", D.J. DiMaria, J. Appl. Phys. 52, 7251 (1981).

18. "Insulator Physics and Engineering: Electrically-Alterable Read- Only-Memories", D.J. DiMaria, Journal de Physique 42, C4-1115 (1981).

19. "Study of Charge Trapping as a Degradation Mechanism in Electrically-Alterable Read-Only Memories", C. Falcony, D.J. DiMaria, D.W. Dong, and K.M. DeMeyer, J. Appl. Phys. 53, 43 (1982).

20. "A New Low-Voltage Si-Compatible Electroluminescent Device", D.J. Robbins, C. Falcony, D.J. DiMaria, D.W. Dong, J.F. DeGelormo, I.F. Chang, and D.B. Dove, IEEE Electron Dev. Lett. EDL-3,147 (1982).

21. "Experimental Test of the Quantum Mechanical Image-Force Theory", A. Hartstein, Z.A. Weinberg and D.J. DiMaria, Phys. Rev. B25, 7174 (1982).

22. "Extended Cyclability in Electrically-Alterable Read-Only Memories", D.J. DiMaria, D.W. Dong, C. Falcony, and S.D. Brorson, IEEE Electron Device Lett. EDL-13, 191 (1982).

23. "Enhanced Electron Injection into SiO_2 Layers Using Granular Metal Films", C. Falcony, D.J. DiMaria, C.R. Guarnieri, J. Appl. Phys. 53, 5347 (1982).

24. "Light Emission from Electron Injector Structures", T.N. Theis, J.R.Kirtley, D.J. DiMaria and D.W. Dong, Phys. Rev. Lett. 50, 750 (1983).

25. "Hot-electron picture of light emission from tunnel junctions", J.R. Kirtley, T.N. Theis, J.C. Tsang, and D.J. DiMaria, Phys. Rev. B, 27, 4601 (1983).

26. "A Study of The Electrical and Luminescence Characteristics of a Novel Si-Based Thin Film Electroluminescent Device, D.J. Robbins, D.J. DiMaria, C. Falcony, and D.W. Dong, to be published J. Appl. Phys. (Aug).

27. "Characterization of Plasma-Enhanced Chemically-Vapor-Deposited Silicon-Rich Silicon Dioxide/Thermal Silicon Dioxide Dual Dielectric Systems", S. Yokoyama, D.W. Dong, D.J. DiMaria, and S.K. Lai, accepted J. Appl. Phys.

28. "Charge Transport and Trapping Phenomena in Off-Stoichiometric SiO_2 Films", D.J. DiMaria, D.W. Dong, C. Falcony, T.N. Theis, J.R. Kirtley, J.C. Tsang, D.R. Young, F.L. Pesavento, and S.D. Brorson, accepted J. Appl. Phys.

29. "Silicon-Rich SiO_2 and Thermal SiO_2 Dual Dielectric for Yield Improvement and High Capacitance", S.K. Lai, D.J. DiMaria and F.F. Fang, accepted IEEE Trans. Elec. Devices.

30.   "Enhanced Conduction and Minimized Charge Trapping in Electrically- Alterable Read-Only-Memories Using Off-Stoichiometric Silicon Dioxide Films", D.J. DiMaria, D.W. Dong, and F.L. Pesavento, submitted J. Appl. Phys.

31.   "Ellipsometry Measurements of Polycrystalline Silicon Films", E.A. Irene and D.W. Dong, J. Electrochem. Soc. 129, 1347, (1982).

32.   "Silicon Oxidation Studies: Measurement of the Diffusion of Oxidant in SiO_2 Films", E.A. Irene, J. Electrochem. Soc., 129, 413 (1982).

IV. PAPERS SUBMITTED WITH THIS REPORT

1.   "Light Emission from Electron Injector Structures", T.N. Theis, J.R. Kirtley, D.J. DiMaria and D.W. Dong.

2.   "Hot-electron picture of light emission from tunnel junctions", J.R. Kirtley, T.N. Theis, J.C. Tsang, and D.J. DiMaria.

3.   "A Study of The Electrical and Luminescence Characteristics of a Novel Si-Based Thin Film Electroluminescent Device", D.J. Robbins, D.J. DiMaria, C. Falcony, and D.W. Dong.

4.   "Characterization of Plasma-Enhanced Chemically-Vapor-Deposited Silicon-Rich Silicon Dioxide/Thermal Silicon Dioxide Dual Dielectric Systems", S. Yokoyama, D.W. Dong, D.J. DiMaria, and S.K. Lai.

5.   "Charge Transport and Trapping Phenomena in Off-Stoichiometric SiO_2 Films", D.J. DiMaria, D.W. Dong, C. Falcony, T.N. Theis, J.R. Kirtley, J.C. Tsang, D.R. Young, F.L. Pesavento, and S.D. Brorson.

6. "Silicon-Rich SiO_2 and Thermal SiO_2 Dual Dielectric for Yield Improvment and High Capacitance", S.K. Lai, D.J. DiMaria and F.F. Fang.

7. "Enhanced Conduction and Minimized Charge Trapping in Electrically-Alterable Read-Only-Memories using Off-Stoichiometric Silicon Dioxide Films", D.J. DiMaria, D.W. Dong, F.L. Pesavento, C. Lam and S.D. Brorson.

8. "Ellipsometry Measurements of Polycrystalline Silicon Films", E.A. Irene and D.W. Dong.

9. "Silicon Oxidation Studies: Measurement of the Diffusion of Oxidant in SiO_2 Films", E.A. Irene.

The first paper by Theis et al. is concerned with light emitted as a result of Surface Plasma Polaritons generated when electrons fall from the conduction band of SiO_2 into the Fermi Level of a metal. This is a new technique for studying the hot electron energy distribution in SiO_2. These results also imply the importance of hot electron injection (by elastic tunneling) as a driving mechanism for luminescence from tunnel junctions. We are interested in pursuing work of this type as a means of gaining insight into the important problem of electrical breakdown in SiO_2 which has never been understood.

The second paper by Kirtley et al. is closely related to the first paper and consists of a study of light emission from Metal-insulator- metal tunnel junctions. The results of this work clearly demonstrates that the light emission results from the hot electrons instead of inelastic tunneling mechanisms. It has been shown that this light results from the hot electrons falling from the conduction band of the SiO_2 into the Fermi level of the metal where they excite surface electromagnetic resonances that are coupled out as external radiation through surface roughness.

This work continues to support our contention that these techniques offer new methods for obtaining valuable new information concerning the behavior of hot electrons in SiO_2. Measurments of this type require the ability to pass relatively large current densities through SiO_2 without destroying the samples. This is possible in our case by the use of the Si-rich charge injector technology. The current densities available with the use of this technology far exceed earlier expectations.

The third paper by Robbins et al. involves a practical use of the Si-rich charge injectors in conjunction with an active luminescent ZnS:Mn layer to build a new novel Electroluminescent Device. It is clear from this work that there are many new device possibilities that can be expected that would use these charge injector structures.

The fourth paper by Yokoyama et al. clearly demonstrates that Si rich SiO_2 can be made using a plasma-enhanced CVD technique as well as the more conventional CVD technique previously used. The relative insensitivity of the first order characteristics of these layers increases our confidence that this technique can be adapted for use in a manufacturing environment. This paper also explores in detail the dependence of the characteristics as a function of composition and demonstrates as expected that an optimum composition exists that is a weak function of the actual structure under consideration.

The fifth paper by DiMaria et al. is concerned with the electronic transport characteristics of Si rich SiO_2 and demonstrates a new application for use in the intervening oxide in Charge Storage Devices (with a lower excess Si concentration) as well as in the Charge Injectors. This application resulted from the surprising observation that the presence of the Si rich material in these intervening layers greatly reduces the electron trapping effects. This has great practical significance since electron trapping in these layers limits the number of useful write-erase cycles for these devices. It has been observed that the use of this technology enables us to increase the number of write-erase cycles from 10^6 to 10^{10} which opens up many new uses for these charge storage devices. The excess silicon present in the off-stoichiometric oxide grows

in clusters, probably < 30 Å in diameter. These silicon islands behave as potential energy wells 3 eV to 4 eV deep from the bottom of the conduction band of the SiO_2. The dominant conduction mechanism is controlled by the tunneling of electrons between the silicon islands.

The sixth paper by Lai et al. describes another use for Si rich SiO_2 layers. This application is concerned with the use of these layers to improve the electrical breakdown characteristics of thin SiO_2 capacitors for application to storage capacitors in dynamic memory devices. Previous work had shown that electron traps near the electrode can shield out effects due to the high electric fields present in the vicinity of asperities and thus result in a tighter breakdown distribution. This paper shows that Si rich SiO_2 layers located adjacent to the metal-insulator interface can also produce this effect. Although some improvment has been observed if the Si-rich layer is located at the anode the greatest improvment is observed when it is at the negative electrode (cathode). One concern for this application has to do with the response time of capacitors using this technology. Some field effect transistor structures have been made to investigate this point. It was observed that the response time is less than 2 nanoseconds and could not be resolved by this measurment. We conclude that the response time does not seem to be an important limitation.

The seventh paper covers material that is closely related to the material covered in the fifth paper.

The eighth paper is concerned with the use of ellipsometry to measure the thickness of polycrystalline silicon films and suggests techniques to make this possible to account for the scattering properties of these films.

Finally, in the ninth paper the diffusion of the oxidant in SiO_2 is studied using an oxidation delay time technique. This is possible due to the *in situ* ellipsometer which makes such measurements feasible.

V. TECHNICAL CONTRIBUTORS TO THIS CONTRACT

C. Aliotta

J. Blum

S.D. Brorson

J. Calise

I.F. Chang

J.F. DeGelormo

K.M. DeMeyer

D.J. DiMaria

D.W. Dong

D.B. Dove

L.M. Ephrath

C. Falcony

F.J. Feigl

B. Ginsberg

F.F. Fang

C.R. Guarnieri

A. Hartstein

E. Irene

J.R. Kirtley

S.K. Lai

A. Lamberti

F.L. Pesavento

W. Reuter

D.J. Robbins

C.M. Serrano

P. Solomon

T. Theis

E. Tierney

J.C. Tsang

Z.A. Weinberg

S. Yokoyama

D.R. Young

ACKNOWLEDGMENTS

We are saddened to report the passing of our colleague - David Dong - who was one of the most prolific contributors to this work. His contributions were in the area of materials studies and materials preperation. He was personally responsible for making most of the Si rich SiO_2 samples and he also made the original suggestion that led to the development of the ultra low trapping oxides. We are going to miss Dave very much both technically and personally. He can never be replaced.

Light Emission from Electron Injector Structures

T. N. Theis, J. R. Kirtley, D. J. DiMaria and D. W. Dong

IBM Thomas. J. Watson Research Center

P. O. Box 218

Yorktown Heights, New York 10598

Surface plasmon polariton mediated luminescence is observed when electrons are injected into thin Al films from the conduction band of SiO_2. These electron injector structures are strikingly similar to light emitting tunnel junctions, although tunneling can be ruled out as the driving mechanism. The emission arises from the energy relaxation of the steady-state hot electron distribution which exists in the metal under continuous current injection. The same mechanism must explain much of the luminescence from tunnel junctions.

Since the seminal papers of Lambe and McCarthy[1,2], it has been generally accepted that light emission from metal-insulator-metal (MIM) tunnel junctions results from a two-stage process. First, an electron tunnels inelastically, losing its energy to a collective excitation of the junction. Second, in the presence of surface roughness, this excitation may radiate. Since the energy loss occurs in the insulating region of the junction,[3,4] inelastic tunneling should most efficiently excite electromagnetic modes with large energy density in this region. Theoretical attention[4,5] was therefore initially focused on the "slow wave" or junction mode,[6] with fields concentrated between the metal electrodes. However, radiation from electrodes consisting of many small metal balls[7-9] has been shown to be mediated by localized plasmons.[9-11] More recently light emission has been demonstrated via the "fast" surface plasmon polariton, which has a maximum energy density at the outer electrode surface.[12-15] To explain the efficiency with which the fast mode is excited, Laks and Mills[16] proposed a phenomenological model in which the inelastic tunneling current fluctuations extend into the metal electrodes on both sides of the insulating junction. However, the excitation efficiency appears to be much higher than predicted by this theory.[14] This and several other puzzling results, discussed at length in Refs. 14 and 17, have suggested the importance of an alternate excitation mechanism, the injection of hot electrons into the metal by *elastic* tunneling.

If this mechanism is correct, it should not matter *how* the electrons are injected, as long as they enter the metal with energies several eV in excess of the Fermi energy. Here we report the observation of surface plasmon polariton mediated light emission when electrons are injected from the conduction band of SiO_2 into a thin Al film. The SiO_2 layer is sufficiently thick (50 nm) that no direct tunneling into the metal, elastic or inelastic, is possible. Nevertheless, the distinguishing features of light emission from tunnel junctions are reproduced. A similar process when the current injection is by elastic tunneling is strongly implied.

The devices used in these experiments are called electron injector structures.[18,19] Figure 1 shows the energy band diagram of such a device under positive bias. It is a metal-

oxide-semiconductor sandwich, with a layer of silicon-rich oxide (Si-rich SiO_2) between the stochiometric oxide (SiO_2) and the degenerately doped n-type Si substrate. The Si-rich material is two phase, containing many tiny ($<\sim$ 5nm) inclusions or islands of Si in an SiO_2 matrix[20]. Conduction through this layer is predominantly by direct tunneling from Si island to Si island. Electrons subsequently enter the conduction band of the stochiometric oxide at thermal energies by Fowler-Nordheim tunneling. Local enhancement of the electric field near each Si island allows the tunneling to proceed at comparatively low bias voltages, yet the islands are packed densely enough that the fields and currents in the bulk of the SiO_2 are laterally uniform[18].

Once in the SiO_2 conduction band, electrons are quickly swept to the opposite interface. Injected into the metal electrode, they rapidly lose energy by electron-electron and electron-phonon interactions on a time scale of $\sim 10^{-15}$ sec. In contrast, the generation of a surface plasmon polariton via an electronic collision with the outer wall of the metal electrode, as shown in Fig. 1, requires $\sim 10^{-12}$ sec.[21] Continuous current injection thus produces in the metal a steady-state distribution of hot electrons with energies depending on both the initial injection energy and the relative strength of the various energy loss channels. The relatively weak plasmon loss will give rise to light emission without greatly affecting the distribution.

In order to unequivocally demonstrate this luminescence mechanism, we have fabricated electron injector structures on periodic corrugated gratings. The gratings were ion milled into Si substrates as described in Ref. 14. Successive layers of Si-rich SiO_2 and stochiometric SiO_2 were then formed by chemical vapor deposition,[18,20] preserving the corrugated pattern as shown in Fig. 2a. Following a $1000°C$-N_2-anneal, circular Al electrodes, ~ 0.017 cm^2 in area and 25 nm thick, were evaporated through a mask. This was followed by a forming gas anneal at $400°C$ for 20 min.

The devices could typically be biased at 20-25 volts with time averaged current densities of about 10^{-2} A/cm^2. As current was passed through the devices, charge was

trapped in the oxide, lowering the field at the Si-rich SiO_2 - SiO_2 interface. The bias voltage, therefore, had to be gradually increased to maintain a constant current. It was possible to pass about 2 coul/cm² through a typical device before sufficiently large fields were built up to cause a destructive breakdown. The emitted light was collected by a spherical mirror, focused at the input of a single pass monochromator, and detected by photon counting techniques. An aperture in front of the collecting mirror allowed angular resolution of the spectra.

Referring to the coordinate system of Fig. 2a, light was collected in the plane defined by $\phi=0$, while θ could be varied. The resonance condition for coupling of a surface plasmon polariton of wave number q to light by a grating of periodicity a then takes the simple form

$$k_{||} \equiv (\omega/c) \sin \theta = q \pm n(2\pi/a) \quad n = 1,2,...., \tag{1}$$

where $k_{||}$ is the wave number parallel to the surface. For a fixed value of θ, Eq. (1) is satisfied only at discrete energies, resulting in characteristic, angle-tunable peaks in the emission spectrum. One such peak is evident in the spectrum in Fig. 2b, obtained at $\theta=13°$ from a device fabricated on a grating of 815 nm period and 36 nm amplitude. The bias voltage was adjusted to maintain a constant current of 100 μA. Also included in the figure is a plot of the specularly reflected light from a collimated white light source directed at the sample at an angle of $\theta=-13°$ and detected in the same geometry as the emitted light. The reflectivity dip corresponds to the satisfaction of the resonance condition, Eq. (1), and the conversion of incident photons to surface plasmon polaritons. The exact match in energy between this dip and the emission peak shows that the same mode is involved in the emission. The peak position is well predicted by Eq. (1) with n=1 only if the emission is mediated by the "fast" surface plasmon polariton with maximum energy density at the Al-air interface and a dispersion which closely follows the light line.[22] Luminescence from tunnel junctions fabricated on corrugated gratings is mediated by the same mode.[12-15]

The emission peak sits on a continuum background. A calculation following the method of Laks and Mills[16] indicates about half of this background is mediated by surface plasmon polaritons. In other words, much of the integrated intensity is not beneath the sharp peak. This is because surface modes are strongly damped on Al, which is a relatively lossy metal compared with, for example, Ag. The remainder of the background is apparently due to luminescence from the oxide layer, as discussed below.

If the periodic grating is replaced by random roughness with transverse correlations of widely varying length, the result will be a broad band emission spectrum. Smooth substrates were roughened by deposition of a 250 nm thick layer of n-degenerate polycrystalline Si. The roughness of this layer was enhanced by thermally growing and then chemically stripping 50 nm of oxide. Injector structures were then fabricated on the substrates as before. In Fig. 3 the spectrum from a charge injector fabricated on a smooth substrate is compared with that from a device fabricated simultaneously on a rough substrate. The electrical characteristics of the devices were virtually identical. The bias voltage was a pulsed square wave, driving a time-averaged current of 10 μA through the sample. The pulse width was 1 msec with a low duty cycle of 1:20 which helped to discriminate against detector dark counts. The spectra have been normalized for the relative throughput of the spectrometer and detection system, with the throughput at 3.5 eV set arbitrarily to 1.

The radiation from the nominally smooth forward biased sample appears to be dominated by luminescence from the bulk of the oxide as reported by others.[23-24] As would be expected in that case, a nearly identical spectrum (not shown) is obtained if the sample is reverse biased so that electrons flow from the metal to the substrate. By contrast, the roughened sample has a dramatically different spectrum in forward (but not in reverse) bias. The increase in brightness with surface roughness is evidence that the additional light emission, the difference spectrum of Fig. 3, is mediated by surface plasmon polaritons. This spectrum is peaked in the

visible and decreases linearly toward a high energy cutoff. It thus strongly resembles the luminescence spectrum of a tunnel junction fabricated on a roughened substrate.[1]

For tunnel junctions the cutoff has been explained as a consequence of the inelastic tunneling model,[1,10] but the difference spectrum of Fig. 3 shows that hot electron injection can also produce this feature. We propose a simple explanation based on the fact that the average injected electron which generates a surface plasmon has already lost energy through other channels. Neglecting any angular dependence of the initial electronic state occupation function, $f(E)$, in the metal, the number of photons emitted with frequency ν within an interval $d\nu$ can be written as

$$L(\nu) = \int_0^\infty f(E)[1-f(E-h\nu)]P(\nu,E)dE, \qquad (2)$$

where $P(\nu,E)$ is a slowly varying function of ν and E involving the density of surface modes, and the surface plasmon polariton excitation and radiation probabilities. We assume $f(E) \approx f_t(E) + \Delta f(E)$, where $f_t(E)$ is the thermal Fermi distribution, and $\Delta f(E) << 1$. If the plasmon loss channel were strong enough to dominate electron-electron and electron-phonon losses, the appropriate hot electron distribution would be a delta-function, $\Delta f(E) \sim \delta(E-E_i)$, at the injection energy, E_i. Since $P(\nu,E)$ depends weakly on energy and since $E_i >> k_b T$, this would result in an abrupt, step-function cutoff, $L(\nu) \sim P(\nu,E)\Theta(E_i-h\nu)$. However, because surface plasmon polariton generation is a weak channel for energy loss it is more accurate to assume a relaxed hot electron distribution $\Delta f(E) \sim \Theta(E_i-E)$ resulting in the linear onset which is in fact observed: $L(\nu) \sim P(\nu,E)(E_i-h\nu)$.

The injection is indeed approximately monoenergetic as assumed above. The minimum energy with which electrons enter the metal corresponds to the potential step, $E_c - E_f$, between the bottom of the SiO_2 conduction band and the Al Fermi level. A value of 3.6 eV was obtained by Solomon and Dimaria[25] from internal photoemission measurements on similarly processed samples. From the cut off at 4.5 eV we deduce that the electrons are injected with

26

average energies about 1 eV above E_c. We emphasize that the cutoff cannot be associated with a lack of small scale roughness for radiative coupling at high energies, since it is observed to move to much higher energies as the field at the SiO_2-Al interface is increased. This is the first direct evidence of strong electric field heating of carriers in the oxide conduction band. Such measurements are of great interest in understanding dielectric breakdown in wide band gap insulators, and a report of initial results is being prepared.[26]

The difference spectrum of Fig. 3 yields a total luminescent quantum efficiency of about 10^{-6} photons/electron, somewhat lower than the best values reported for tunnel junctions with Ag or Au electrodes.[1,2,8,9] However, the radiative efficiencies of surface plasmon polaritons are much higher for Ag or Au than for the Al electrodes of our injectors. The model of Laks and Mills[5,16], which gives a very good account of the radiative properties of MIM tunnel junctions,[14] predicts that injectors fabricated with Ag electrodes should be about 100 times brighter. The extrapolated efficiency of 10^{-4} is then easily comparable to the best tunnel junction values.

We have demonstrated that light emission from the metal electrodes of suitably roughened or patterned electron injector structures is mediated by surface plasmon polaritons. The luminescence is both driven by and indirectly probes the steady-state distribution of hot electrons excited in the metal and, through the cutoff energy, provides information about the energy distribution of carriers in the injecting oxide conduction band. We have emphasized that these results imply the importance of hot electron injection (by elastic tunneling) as a driving mechanism for luminescence from tunnel junctions. However, another important application of our results may be in the study of conduction processes in wide band gap insulators.

in this work, and C.F. Aliotta for characterising the samples with scanning electron microscopy. We thank A.B. Fowler, D. L. Mills, B. Laks and F. Stern for helpful discussions.

REFERENCES

1. J. Lambe and S.L. McCarthy, Phys. Rev. Lett. 37, 923 (1976).

2. S.L. McCarthy and J. Lambe, Appl. Phys. Lett. 30, 427 (1977); 33, 858 (1978).

3. K.L. Ngai, E.N. Economou, and Morrel H. Cohen, Phys. Rev. Lett. 22, 1375 (1969).

4. L.C. Davis, Phys. Rev. B 16, 2482 (1977).

5. B. Laks and D.L. Mills, Phys. Rev. B 20, 4962 (1979); 21, 5175 (1980).

6. J.C. Swihart, J. Appl. Phys. 32, 461 (1961).

7. S.L. McCarthy and John Lambe, Appl. Phys. Lett. 33, 858 (1978)

8. P.K. Hansma and H.P. Broida, Appl. Phys. Lett. 32, 545 (1978).

9. A. Adams, J.C. Wyss and P.K. Hansma, Phys. Rev. Lett. 42, 912 (1979).

10. Daniel Hone, B. Muhlschlegel, and D.J. Scalopino, Appl. Phys. Lett. 33, 203 (1978).

11. R.W. Rendell and D.J. Scalapino, Phys. Rev. B 24, 3276 (1981).

12. J.R. Kirtley, T.N. Theis and J.C. Tsang, Appl. Phys. Lett. 37, 435 (1980).

13. N.Kroo, Zs. Szentirmay and J. Felszerfalvi, Phys. Lett. 81A, 399 (1981).

14. John Kirtley, T.N. Theis, and J.C. Tsang, Phys. Rev. B 24, 5650 (1981).

15. N. Kroo, Zs. Szentermay and J. Felszerfalvi, Phys. Lett. 86A, 445 (1981).

16. Bernardo Laks and D.L. Mills, Phys. Rev. B 22, 5723 (1980).

17. J.R. Kirtley, T.N. Theis, J.C. Tsang, and D.J.DiMaria, submitted to Phys. Rev. B..

18.   D.J. DiMaria, R. Ghez, and D.W. Dong, J. Appl. Phys. $\underline{51}$, 4830 (1980).

19.   D.J. DiMaria, K.M. DeMeyer, C.M. Serrano, and D.W. Dong, J. Appl. Phys. $\underline{52}$, 4825 (1981).

20.   D.J. DiMaria in The Physics of MOS Insulators, ed. by G. Lucovsky, S.T. Pantelides, and F.L. Galeener (Pergamon, New York, 1980), pp. 1-18.

21.   J.R. Kirtley, unpublished.

22.   For a complete discussion of surface plasmon modes in various planar thin film geometries see E.N. Economou, Phys. Rev. $\underline{182}$, 539 (1969).

23.   P. Solomon and N. Klein, J. Appl. Phys. $\underline{47}$, 1023 (1976).

24.   C. Falcony, Doctoral Dissertation, Lehigh University, 1980, unpublished.

25.   P.M. Solomon and D.J. DiMaria, J. Appl. Phys. $\underline{52}$, 5867 (1981).

26.   T.N. Theis, J.R. Kirtley, D.J. DiMaria, and D.W. Dong, unpublished
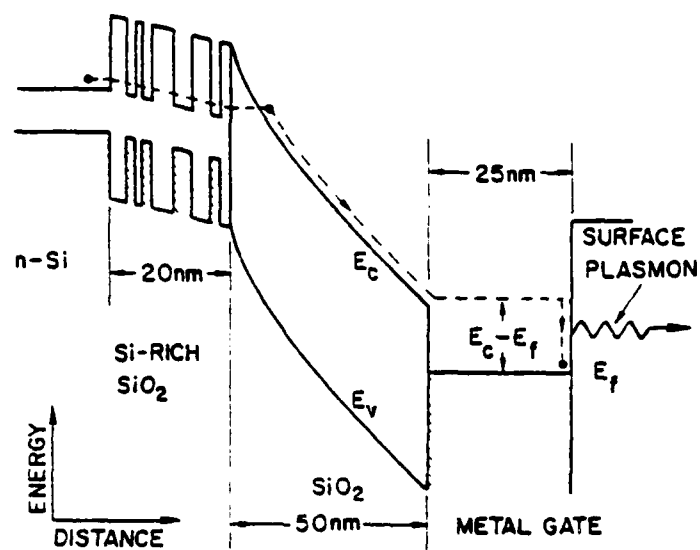
Fig. 1 Energy band diagram of a positively biased electron injector structure.

(a)

(b)
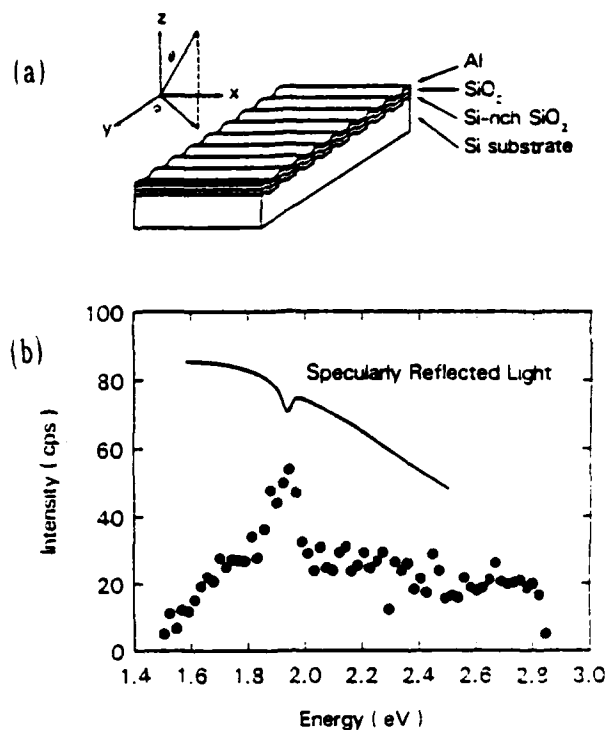
Fig. 2. (a) Schematic drawing of an electron injector fabricated on a substrate with a corrugated grating surface. Luminescence spectra were obtained in the $\phi=0$ plane as a function θ. (b) Luminescence spectrum from the sample at $\theta=13°$ (dots). The sharp emission peak at 1.95 eV corresponds well to the surface plasmon polariton dip in the specular reflection for light incident at $\theta=-13°$, shown by the solid curve.

Fig. 3 Luminescence spectra from rough and smooth electron injector structures. The difference spectrum represents the surface plasmon polariton mediated emission from the roughened sample.

Hot-electron picture of light emission from tunnel junctions *

J. R. Kirtley, T. N. Theis, J. C. Tsang, and D. J. DiMaria

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 8 November 1982)

Metal-insulator-metal tunnel junctions emit optical radiation when biased at voltages in the range 2—4 V. We argue that a complete pict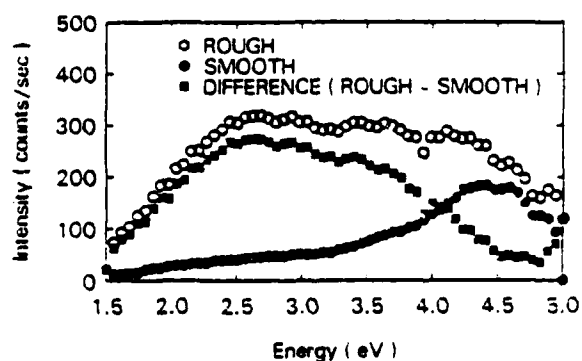ure of this radiation process includes hot electrons, which excite surface electromagnetic resonances, which are in turn coupled to external radiation through surface roughness. This picture is supported by measurements of the temperature and second-metal-electrode thickness dependence of the emission intensities, and by light emission from surface plasmons excited by optical pumping and by charge injection.

I. INTRODUCTION

Lambe and McCarthy[1] pointed out that tunnel junctions emitted optical radiation when biased at voltages V_0 in the range 2—4 V. The emitted radiation became more intense if the junctions were intentionally roughened. The emitted radiation from randomly roughened junctions was broadband with a characteristic linear onset below a critical frequency v_c, related to the junction bias voltage by $hv_c = eV_0$. Light-emitting tunnel junctions have some of the attributes of an attractive display device: They are flat, operate at room temperature in air, require low dc bias voltages, have emission frequencies with an upper cutoff tunable through the visible, and are inexpensive to make. However, the best demonstrated external quantum efficiencies (the number of photons out divided by the number of electrons crossing the tunneling barrier which are approximately the external power efficiencies) are of order 10^{-4}.[2-7] It is, therefore, of practical as well as fundamental interest to understand the emission processes.

Several papers have described light emission from tunnel junctions in terms of a two-step process: (1) inelastically tunneling electrons excite collective electromagnetic oscillations of the junction, and (2) the electromagnetic oscillations emit external radiation.[1,8-11] There are two classes of electromagnetic oscillations relevant to the tunneling junction geometry.[12] The first is the junction or slow mode. This mode has fields and energies located primarily in the junction region, and can be described as primarily electrostatic. Because of the screening between the two metal-insulator interfaces, the junction mode has a speed of propagation much slower than that of light in free space. The second relevant mode is the Ag-air interface fast surface-plasmon polariton. It has fields and charges localized pri-

marily at the Ag-air interface, and propagates at a speed quite close to that of light in free space. Surface roughness is required for either mode to radiate light, since energy and momentum cannot be simultaneously conserved in a radiative transition without roughness scattering.

It has also been reported that radiation from tunnel junctions is dominated by the junction mode.[10,11] In this view the coupling to surface plasmons is relatively efficient, since the slow mode has large field strengths in the tunneling region, but that the radiative step is relatively inefficient. The inefficiency of the radiative step could be attributed to the large difference between the wavelengths of the junction modes and light. Roughness on a scale of 1—10 nm would be required to efficiently couple the junction mode to light. Roughness on this scale, while possibly present, is difficult to characterize to test these ideas experimentally.

By fabricating tunnel junctions on holographically produced gratings with single Fourier components of roughness of order 800 nm, we were able to show that radiation from the fast mode was an important mechanism for light emission from tunnel junctions in general, and, in fact, dominated over radiation from the junction mode in our samples.[6,7]

Since the electric fields associated with the fast surface mode extend throughout the metal-insulator-metal structure, and are strongest at the Ag-air interface, the coupling between the tunneling electrons and the surface plasmons may well occur outside the tunneling barrier region. We therefore make a distinction between inelastic tunneling and hot-electron coupling to surface plasmons: We describe a process in which the electrons lose energy in the barrier region as inelastic. Conversely, we describe a process in which the electrons first tunnel into one of the electrodes before losing energy as hot. A complete description of the tunneling-

electron—surface-plasmon coupling process would include both events as appropriate limits. While such a comprehensive description does not at present exist, important qualitative conclusions about the dynamics of the system result if we make this distinction.

As noted above, light emission from randomly roughened tunnel junctions has a linear onset below a critical frequency. This can be explained quite naturally in the inelastic tunneling picture in terms of the linearly increasing final density of states available to the tunneling electrons with decreasing energy loss.[8] However, it is difficult to make this qualitative explanation fit the data quantitatively. As will be shown below, when the energy dependence of the tunneling matrix elements are taken into account, the predicted onset below the critical frequency is slower than linear, and is, in fact, slower than observed experimentally from tunnel junctions on gratings.

Further, we will present in this paper a number of experimental results which are difficult to explain using the inelastic tunneling picture: (1) The emission from junctions on gratings becomes exponentially less intense as the second-metal electrode is made thicker. The characteristic intensity falloff length does not match the optical screening length, and is dependent on sample-preparation procedures; (2) the emission-peak intensities increase dramatically as the junction temperatures decrease, even though the junction current-voltage characteristics and optical properties are not strongly temperature dependent; (3) similar external efficiencies and radiation patterns result if the surface plasmons are pumped by optical radiation or by charge injection. In neither the optical nor the charge-injection case is inelastic tunneling present.

All of these results can be understood using a three-step or hot-electron model: (1) Hot electrons are injected into the junction structure; (2) the injected hot-electron distribution relaxes primarily through the emission of phonons and cooler electrons, but also through the emission of surface plasmons; (3) the surface plasmons couple out to external radiation through surface roughness. The hot-electron model has important qualitative consequences which will be explored.

II. TUNNELING RESULTS

We have described light emission from tunnel junctions on gratings previously.[6,7] In these papers the results were analyzed in terms of the inelastic tunneling picture. We will show in this paper how the old results and new results to be presented below are more consistent if viewed in the hot-electron picture. We will merely outline previously presented

experimental details here. Al-Al$_2$O$_3$-Ag tunnel junctions were fabricated on (80—1200)-nm periodicity, (0—100)-nm amplitude and holographically produced corrugated gratings. In some samples the photoresist was itself used as a substrate, while in other samples the pattern was transferred from the resist to a Si substrate by ion milling. Tunnel junctions were formed on the grating substrates. The junctions were formed by evaporating 2-mm-wide, 10-mm-long, and 40-nm-thick Al films through mechanical masks, oxidizing the films, and then completing the junctions with a long 2-mm-wide (15—80)-nm-thick Ag film. All electrical measurements were made with a four-terminal technique. The substrates were mounted on the cold finger of a closed-cycle refrigerator with optical windows, and run in vacuum. The emitted radiation was frequency analyzed using a single-pass monochromator and photon counting. Only the light through a narrow aperture 1.4° wide and 9.5° high, in the plane defined by the junction normal and the grating periodicity wave vector, was allowed to enter the spectrometer.

The emitted light was composed of narrow angle-tunable peaks superimposed on a broad background. Analysis of the peak energies as a function of emission angle showed that the radiation was dominated by emission from the Ag-air interface fast surface-plasmon polariton. It is possible that the broad background we observed had contributions from the junction or localized plasmon modes. All of the data presented here has had the broad emission subtracted from the fast-mode peaks. This allowed us to anlyze only the radiation from well-characterized modes.

Our experimentally measured dispersion curves, linewidths, and dependence of peak intensities on grating amplitudes agreed well with a theory of Laks and Mills[13] using the inelastic tunneling model. This should not be surprising since these properties depend only on the electrodynamics of the system. However, there were serious discrepancies between theory and experiment for those properties which depended on the coupling between tunneling electrons and surface plasmons. For example, the observed intensities were at least 35 times stronger than predicted theoretically.[6]

One of the remarkable points of agreement between the theory of Laks and Mills[13] and our experiments occurred when we compared the integrated peak intensities for a single sample for a series of bias voltages. Laks and Mills,[13] following a suggestion by Hone et al.,[8] wrote the radiation from a tunnel junction as

$$\frac{dP}{d\Omega\,d\omega} = |I(\omega)|^2 A(\omega,d,\epsilon),$$ (1)

where $dP/d\Omega\, d\omega$ is the power radiated per unit
solid angle, per unit frequency interval, $A(\omega,d,\epsilon)$ is
an "antenna factor" which depends on the dielectric
properties of the junction structure, but does not de-
pend on the bias conditions, and $|I(\omega)|^2$ is the
power-spectral density of current fluctuations across
the junction, which does depend on the details of the
charge transport across the junction. We define the
operator $\hat{I}$ for current across the junction as

$$\hat{I} = \frac{ie}{\hbar} \sum_{k,q,s} (T_{kq} c_{ks}^\dagger c_{qs} + T_{kq}^\dagger c_{ks} c_{qs}^\dagger) , \qquad (2)$$

where T_{kq} is the tunneling matrix element from a
state k in one metal electrode to a state q in the oth-
er, and $c_{ks}^\dagger$ and c_{ks} are the electron annihilation and
creation operators for a state with momentum k and
spin s in one metal electrode. The power-spectral
density of current fluctuations across the junction
with frequency ω is defined by

$$|I(\omega)|^2 = \sum_f |\langle f|\hat{I}|0\rangle|^2 \delta(\omega - (E_F - E_0)/\hbar) . \qquad (3)$$

This expression has been evaluated in the limit of
small frequencies by Hone et al.[8]:

$$|I(\omega)|^2 = \frac{eI_0}{2\pi} \left[1 - \frac{\hbar\omega}{eV} \right] , \qquad (4)$$

where I_0 is the current per unit area through the
junction. When this simple expression is used to
divide the current fluctuation factor from the ob-
served light-emission intensities from tunneling
junctions on gratings, the reduction is very good.[6]
That is, the antenna factor is relatively independent
of bias conditions for a given junction, as it should
be. This agreement is illustrated in Fig. 1 for an
Al-Al$_2$O$_3$-Ag tunnel junction on an 800-nm periodi-
city photoresist grating for biases from 1.8–2.6 V.

The solid curve in Fig. 1 was obtained by fitting
the logarithm of the experimental values to a
fourth-order polynomial. If we define a reduced χ^2
value for the overlap of the curves as

$$\chi^2 = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{Y(i) - Y_{fit}(i)}{Y(i)} \right]^2 , \qquad (5)$$

we obtain a χ^2 value of 0.0275.

However, the current fluctuation spectrum looks a
good deal different if one takes into account the
finite-energy losses present (2–3 eV) under experi-
mental conditions. A physical interpretation of Eq.
(3) is that electrons are transferred from a high-
energy state in one electrode to a low-energy state in
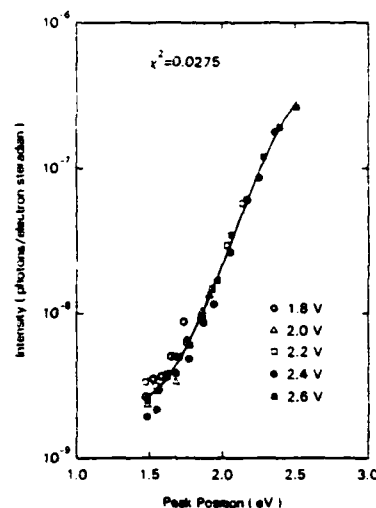the other, and back again. Since more energetic

FIG. 1. Plot of the antenna factor for an Al-Al$_2$O$_3$-Ag
tunnel junction on a grating obtained by dividing the in-
tegrated peak intensities by the expression for the power-
spectral density of current fluctuations across the tunnel
junction derived by Hone et al. [Eq. (4)]. The antenna
factor should be independent of the bias conditions, so the
curves for different bias voltages should overlap, as they
do. The scatter between the experimental values and the
smooth polynomial curve fitted to them has a χ^2 value of
0.0275.

electrons are more likely to penetrate the tunneling
barrier, low-frequency fluctuations, in which the
electrons on the average have more energy, are more
likely to occur than high-frequency fluctuations. A
simple technique for accounting for the dependence
of the barrier penetration probabilities on energy is
to split the tunneling matrix element $|T_{kq}|^2$ into
two parts: One which transfers the electron from
the left to the right electrode at the initial energy,
and a second which transfers the electron from the
right electrode to the left at the initial energy minus
the energy associated with the fluctuation frequency.
This is clearly an approximation to the real process,
but is probably more correct than simply taking the
low-frequency limit to the barrier penetration proba-
bilities.

If we take the "KB approximation[14] for the elec-
tronic wave functions, the low-temperature limit for
the Fermi occupation functions, and use the nota-
tion and approximations that Simmons[15] uses for
the case of elastic tunneling, the power-spectral den-
sity of current fluctuations across the tunneling bar-
rier per unit area is given by

$$|I(\omega)|^2 = \frac{2me^2}{h^3}\left[(eV-\hbar\omega)\int_0^{\epsilon_F-(eV-\hbar\omega)}dE_x\exp\left[-\int_0^l dx\,[K(x)+Q(x)]\right]\right.$$

$$\left.+\int_{\epsilon_F-(eV-\hbar\omega)}^{\epsilon_F}dE_x(\epsilon_F-E_x)\exp\left[-\int_0^l dx\,[K(x)+Q(x)]\right]\right],\qquad(6)$$

where ϵ_F is the Fermi energy of one of the metal electrodes, V is the voltage bias across the junction,

$$K(x)\equiv\left|\left[\frac{2m}{\hbar^2}\right][U(x,V)-E_x]\right|^{1/2},$$

and

$$Q(x)\equiv\left|\left[\frac{2m}{\hbar^2}\right][U(x,V)+\hbar\omega-E_x]\right|^{1/2}.$$

$U(x,V)$ is the position- and voltage-dependent barrier potential seen by the tunneling electrons, x is the Cartesian coordinate normal to the junction interfaces, l is the thickness of the barrier, and E_x is the kinetic energy of the tunneling electrons normal to the interface.

A comparison of the predictions of Eqs. (4) and (6) are plotted in Fig. 2 for an Al-Al$_2$O$_3$-Ag junction with $\phi_1=2.5$ eV, $\phi_2=3.5$ eV, and $l=1.3$ nm. The curves are normalized to the zero-frequency power-spectral density $|I(0)|^2=eI_0/2\pi$. The solid curves are the results of the full expression [Eq. (6)]; the

dashed curves are the low-frequency limit [Eq. (4)]. As expected, the high-frequency fluctuations fall below the curves predicted by the simple expression. As can be seen in Fig. 3, when the power-spectral-density term of Eq. (6) is divided out of the experimental data to obtain the antenna factor, the overlap between the curves for the different bias voltages is appreciably worse (a χ^2 of 0.0637) than that obtained using the simple expression (Fig. 1): The simple expression works better than it should. We argue that this is because the physics involved in the emission process is very different from the inelastic tunneling model, and that the good agreement of the simple inelastic model is probably fortuitous.

Another discrepancy was observed when the emission-peak energies were measured as a function of Ag-film thickness. These results have been detailed and supported by numerical results in Ref. 6. We outline them here to emphasize the close connection with the other results discussed in this paper. The observed peak widths and energies did not

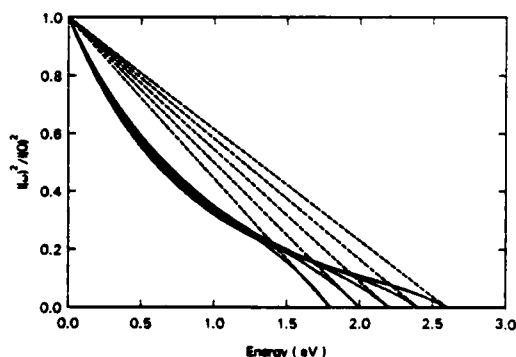FIG. 2. Plot of the power-spectral density of current fluctuations for a tunnel junction with a trapezoidal barrier with heights $\phi_1=2.4$ eV, $\phi_2=3.5$ eV, and thickness $l=1.3$ nm. The dashed curves correspond to the predictions of the simple expression of Hone et al. [Eq. (4)]; the solid curves result from accounting for the energy dependence of the barrier penetration probabilities [Eq. (6)]. The more complex expression falls below the simple one, indicating that one might expect onsets in intensity below the critical frequencies that are slower than linear.

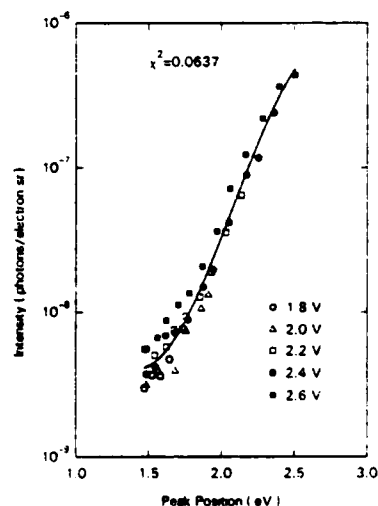FIG. 3. Plot of the data of Fig. 1 for the antenna factor using the power-spectral density of fluctuations across the tunnel junction of Eq. (6). The curves for the different bias voltages do not overlap as well as in Fig. 1: Accounting for the energy dependence of the barrier penetration probability makes the agreement between the inelastic tunneling model and experiment worse, raising the χ^2 value to 0.0637.

change appreciably as the Ag was made thicker, but the emission efficiencies (normalizing out the bias currents) became exponentially smaller. Figure 4(a) shows a plot of the relative emission efficiencies (photons/electron sr) for a set of Al-Al$_2$O$_3$-Ag junctions on a 1.1-μm Si grating. The emission intensities fell off with Ag-film thickness d with the approximate relation $I = Ce^{-d/t}$, where t was about 20 nm. There may also have been some weak energy dependence: The peaks at lower energies appeared to fall off more rapidly with Ag thickness than those at higher energies. Also plotted in Fig. 4(a) is the relative optical screening measured from Ag films on control quartz wafers, using the attenuation of the 632.8-nm line from a helium-neon laser. This screening length does not depend strongly on energy in this region. The experimentally determined falloff length was significantly longer than the optical screening length for this sample. In contrast, Fig. 4(b) shows the dependence of emission-peak efficiencies on Ag-film thickness for a set of junctions on an 815-nm periodicity grating. In this case the characteristic falloff lengths were shorter than the optical screening length.

The radiative efficiency of surface plasmons is relatively insensitive to Ag-film thickness, as has been demonstrated by Moreland, Adams, and Hansma.[16] Therefore, the dependence of emission intensities on film thickness that we observe must reflect the efficiency of production of surface plasmons by tunneling electrons. The theory of Laks and Mills for light emission from tunnel junctions on grat-

ings[13] includes a current fluctuation driving term that models the spatial and frequency dependences of inelastic tunneling transitions in the junction. If the inelastic transitions occur only in the tunneling barrier region, the falloff in emission intensities with Ag-film thickness should be given by the optical screening length, since the surface-plasmon fields decay from the Ag surface with this characteristic length. If, however, the inelastic transitions extend throughout the junction volume, there should be no attenuation with Ag-film thickness. In neither case should coupling of the tunneling electrons to surface plasmons in the bulk of the junction result in an attenuation length shorter than the optical screening length. We, therefore, interpret the results of Fig. 4 to mean that the coupling of the tunneling electrons to surface plasmons occurs not in the junction volume, but primarily at the Ag-air interface. The results of Fig. 4 are then a measure of the hot-electron attenuation length in the Ag films.

The attenuation lengths for the two sets of junctions differ because of differences in sample fabrication: The first difference is that the samples of Fig. 4(a) were made on an etched Si wafer while those of Fig. 4(b) were made on a photoresist grating. Electron micrographs show that the Si wafer is significantly smoother. The small-scale roughness present in the photoresist films could act as agglomeration sites, tending to reduce the average film grain sizes. The second difference is that the Ag films in Fig. 4(a) were evaporated at 1–2 nm/sec while those of Fig. 4(b) were evaporated at 0.2–0.4 nm/sec. The slower evaporation rate might be expected to result in higher defect densities. The differences in substrate roughness and evaporation rates could both be expected to reduce the hot-electron mean free path for the sample of Fig. 4(b) relative to that of Fig. 4(a). The value of ~20 nm for the hot-electron mean free path at 2 eV for the sample of Fig. 4(a) agrees well with 18 nm reported by Crowell and Sze.[17]

It is reasonable to expect hot electrons to couple to surface plasmons more strongly at the Ag-air interface than in the bulk of the film. The argument goes as follows: The wavelengths for electrons at the Fermi surface of a free-electron metal are of order 0.1 nm. The surface-plasmon fields decay into the metal with an attenuation length of about 10 nm. Therefore, the overlap integrals involved in the hot-electron–surface-plasmon matrix elements due to the bulk of the metal tend to be small. On the other hand, the surface-plasmon fields at the Ag-air interface change rapidly on a scale of 0.1 nm, and the surface contributions to the matrix elements can be large. Another way of saying this is that (in the free-electron model) momentum cannot be conserved
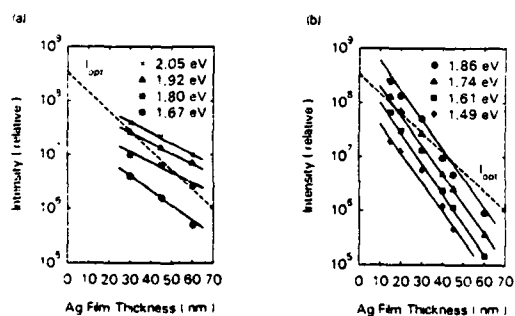
FIG. 4. Comparison of the Ag-film thickness dependence of emission intensities for Al-Al$_2$O$_3$-Ag light-emitting tunnel junctions on gratings. Junctions fabricated on a relatively smooth Si substrate and with fast Ag evaporation rates have intensities that decrease with Ag-film thickness at a slower rate than junctions fabricated on photoresist substrates with slow Ag evaporations. In neither case is the characteristic length the optical screening length, as one would expect if the interaction between the tunneling electrons and the surface plasmons occurred in the barrier region.

in surface-plasmon emission from hot electrons in the bulk metal, but the surface can provide the required momentum. Large surface contributions have, in fact, been observed in photoemission from metals.[18] We have developed a simple theory for surface coupling of tunneling electrons to surface plasmons which will be published separately.[19] The important point for this discussion is that the dependence of emission efficiencies on Ag-film thickness in such a model does not follow the optical attenuation lengths. Therefore, the interaction between tunneling electrons and surface plasmons cannot be exclusively in the barrier region.

A second indication of hot-electron effects comes from measurements of the temperature dependence of the light emission from junctions on gratings. In Fig. 5(a) we plot the peak intensity from an Al-Al$_2$O$_3$-Ag junction on a 10-nm amplitude, 1225-nm periodicity grating as a function of temperature. The sample was rotated with respect to the optical
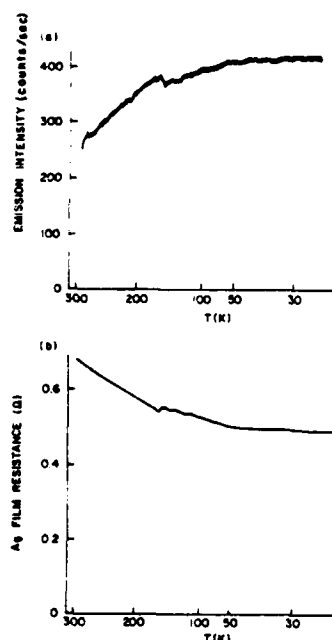

FIG. 5. Emission intensity vs temperature and Ag-film resistance vs temperature for an Al-Al$_2$O$_3$-Ag tunnel junction on a 1225-nm periodicity, 10-nm amplitude quartz grating substrate. The sample was rotated so that the surface-plasmon emission peak was observed at 1.77 eV. The junction bias voltage was held constant at 2.13 V with a nearly constant bias current. As the temperature was lowered the emission intensity increased, presumably because of an increase in the hot-electron mean free path in the Ag, as evidenced by the decrease in the Ag-film resistance.

aperture until the surface-plasmon emission peak was at 1.77 eV. The sample bias current was held constant at 44.1 mA; the sample bias voltage was approximately 2.13 V. The temperature was swept down over a period of about 90 min. The junction resistance changed slightly as the temperature was lowered such that the voltage across the junction increased by 5% in going from room temperature to 25 K. This slight change in the bias voltage was not nearly large enough to explain the ~60% increase in the emission intensity as the sample temperature was lowered. In addition, the emission-peak widths did not change appreciably as the temperature was lowered from room temperature to 25 K. This indicated that the damping of the surface plasmons was not strongly dependent on temperature. We determined that the optical attenuation lengths of the Ag films were not changing with temperature by passing the beam from a He-Ne laser through a 28-nm Ag film on a glass substrate and measuring the transmitted intensity. The optical transmisssion of the film at 632.8 nm was unchanged to within our experimental accuracy of 10% over a temperature range of 300−25 K.

In Fig. 5(b) we plot the Ag-counterelectrode film resistance as a function of temperature for the same sample as that of Fig. 5(a). The Ag-film resistance went down about 20% as the temperature decreased. The resistance of the film as a function of temperature had sharp structure, presumably due to morphological changes. This structure was quite reproducible but hysteretic: The changes occurred at higher temperatures when the temperature was swept up than when it was swept down. The salient point is that the structure in the Ag-film resistance was reproduced in the junction emission intensity. The changes in the strip resistance were too small to affect the bias of the junction significantly: The effective junction resistance was about 50 Ω at this bias voltage; the Ag-strip resistance was less than 1 Ω and changed less than 0.2 Ω. We therefore interpret these results to mean that in addition to the lowering of the Ag-film resistance due to a freezing of the metal phonons, there were also electron scattering mechanisms in the Ag film that changed as the temperature was lowered. The hot-electron mean free path increased as the temperature was lowered, increasing the probability for the hot electrons to get to the Ag-air interface to couple to surface plasmons, and thereby increasing the emission efficiency. These experimental results are difficult to explain in an inelastic tunneling model, since it is difficult to understand how a change in the scattering properties of the Ag film could affect the emission efficiency if the tunneling-electron—surface-plasmon interaction was in the tunneling region.

III. OPTICAL PUMPING

Further evidence in favor of hot-electron effects in light emission from tunnel junctions came from experiments in which the surface plasmons were pumped by other means. Laser radiation of a set energy incident on a tunnel junction on a grating exhibits a sharp reflectivity minimum at an angle corresponding to the angle at which the laser couples to surface-plasmon polaritons through the grating roughness. The decay of these surface plasmons must result in a distribution of hot electrons in the metal with energies up to the incident laser energy. The hot electrons can subsequently excite surface plasmons of lower energy, and plasma-mediated radiation at lower energies is indeed observed.[20,21]

The laser- and tunneling-electron-pumped data show striking similarities. Figure 6 shows a plot of the Stokes-shifted radiation from a tunnel junction irradiated with 60 mw of laser radiation at 2.41 eV. The incident laser beam was held at the resonance angle so that about 90% of the incident energy was taken out of the specular beam to excite the surface wave. The scattered light was apertured as described above and frequency-analyzed using a double monochromater. The same optical system was used for both the laser-pumped and tunneling-pumped emission described in this section. The junction of Fig. 6 had a molecular monolayer of 4-nitrobenzoic acid absorbed on the aluminum oxide in the Al-Al$_2$O$_3$-Ag structure, and therefore showed sharp surface-enhanced Raman scattering peaks in addition to broader emission peaks corresponding to coupling out of the Ag-air fast surface-plasmon po-

lariton. The broad emission peaks moved in frequency and changed in intensity as the observation angle was changed, just as the emission peaks from tunneling pumping did. Figure 6 shows the emission spectrum of the same sample, with the same collection optics, but without an incident laser beam and with the junction biased at 2.41 V and 25 mA. The absence of sharp Raman scattering peaks in the tunneling-pumped emission spectrum is to be expected because the injected electron-energy distribution was relatively broad, but the broad emission peaks are present, regardless of how (optical pumping or tunnel current injection) the sample is excited.

Figure 7 contains a comparison of the plasmon emission-peak intensity as a function of peak energy (obtained simply by rotating the observation angle) for tunneling and optical pumping, with the total pumping power per unit area imaged on the spectrometer slit normalized out. Not only does the optically pumped data show a linear onset below the pump energy, but the total external quantum efficiency is very similar to that for the tunneling pumped case. The linear onset of emission intensities from light-emitting tunnel junctions below the bias voltage energy has been used as an argument for an inelastic tunneling process. The fact that this linear onset also occurs, with comparable efficiency, in the optical pumping case in which there are no tunneling electrons, indicates a hot-electron light-emission mechanism.

We argue that similar processes are occurring in the two cases: Hot-electron distributions are introduced by tunnel injection or optical absorption, the electron distributions relax in energy partially
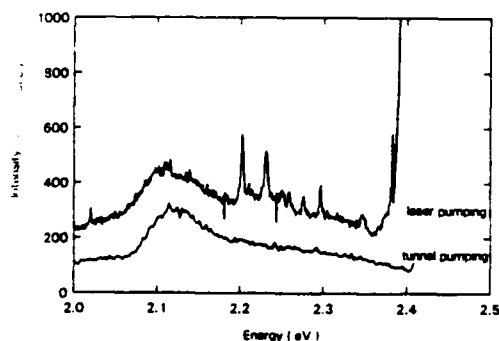
FIG. 6. Comparison of the surface-plasmon emission from an Al-Al$_2$O$_3$-Ag tunnel junction for laser pumping and tunneling pumping. The sharp lines in the laser-pumped case are surface-enhanced Raman-scattered light from a monolayer of 4-nitrobenzoic acid included in the junction region. The broad peaks at 2.12 eV are the surface-plasmon peaks.
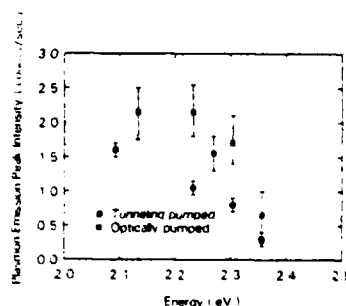

FIG. 7. Comparison between the plasmon emission-peak intensities from an Al-Al$_2$O$_3$-Ag tunnel junction under optical and tunneling pumping. The similar linear onset below the pump energy and similar total external quantum efficiencies indicate that similar processes are occurring in the two cases.

through the emission of surface plasmons, and the surface plasmons couple out to light through the grating roughness.

IV. CHARGE-INJECTION STRUCTURES

A third method for producing surface-plasmon emission from a thin metal film is by charge injection.[22] A schematic of the charge-injection structures[23,24] we used is shown in Fig. 8. A single-crystal degenerately n-type—doped (0.001 Ω cm) Si wafer was used as a substrate. A Si-rich ($\sim 15\%$ excess Si) SiO_2 layer (20—50)-nm thick was deposited on the substrate by chemical vapor deposition. This was followed by a stoichiometric SiO_2 layer of about 50 nm and then a gate-metal electrode of typically 20—25 nm was deposited by evaporation. The concentration of Si in the SiO_x was determined from the ratio of silane (SiH_4) to nitrous oxide (N_2O) in the gas phase during the chemical vapor deposition.[23,24] The Si-rich SiO_2 had small ($< \sim 5$-nm-diam) Si inclusions.[25] Electrons tunnel from Si inclusion to Si inclusion, until they reach the interface with the SiO_2 layer, where they are injected into the SiO_2 conduction band through field-enhanced Fowler-Nordheim tunneling. The electric field enhancement is caused by the nonplanar shape of the Si islands. This lowers the gate voltage required for current flow. Once in the conduction band, the electrons proceed rapidly through the SiO_2 and are injected into the gate-metal electrode. The principle advantage of the charge-injection structure over a planar Si-SiO_2-M structure is that it can inject large current densities ($> 10^2$ A/cm^2) for long periods of time at moderate gate voltages without destructive

FIG. 8. Schematic energy-band diagram of a Si-Si-rich SiO_2-SiO_2-M charge-injection structure. The electrons tunnel through the Si-rich SiO_2 into the conduction band of the SiO_2, are injected into the gate-metal film at an energy corresponding to the step between the bottom of the SiO_2 conduction band and the Fermi level of the metal, and produce surface-plasmon radiation.

breakdown of the SiO_2 layer.

These structures had current-voltage characteristics that approximately followed the Fowler-Nordheim relation $I = AV^2e^{-B/V}$. We typically ran at biases of 25 V with up to 100 μA current through a sample area of 1 mm^2 (10^{-2} A/cm^2). As current was passed through the devices, charge was trapped in the oxide, lowering the electric field at the injecting Si-rich SiO_2-SiO_2 interface so that more voltage was required to keep the current through the device constant. Eventually, sufficient charge was trapped and sufficiently large fields were built up to cause a destructive breakdown. We found that we could pass about 2 C/cm^2 through the devices before they failed.

The charge-injection devices had some intrinsic roughness because of the Si inclusions in the Si-rich oxide. We introduced additional roughness to some of the samples by depositing polysilicon on the Si substrates, doping it n-type degenerate with POCl$_3$, oxidizing the polysilicon to an SiO_2 thickness of about 50 nm, and then stripping the oxide. The oxide grows preferentially faster along grain boundaries and certain crystallite orientations, leaving a rough surface upon removal. Controlled roughness was also produced in some samples by etching holographically produced gratings into the Si substrates before the deposition of the SiO_2.

Figure 9 compares the emitted radiation from a polysilicon-roughened Al-gate charge-injection structure (LUM24-H) with that from an identical structure laid down on a single-crystal Si substrate (LUM24-X1). The Al-gate electrodes were 25 nm thick and were annealed at 400°C for 30 min in forming gas after deposition. The devices were pulsed with the gate at about 20—25 V positive with respect to the substrate, with an average current through the sample of 10 μA. The pulses were 1-msec wide with a 5% duty cycle. The low duty cycle helped to discriminate against phototube dark counts. The spectra in Fig. 9 are normalized for the relative spectrometer throughput, with the throughput at 400 nm set arbitrarily to 1. The emission was collected with $F/4.8$ mirror collection optics. The radiation from the nominally smooth sample appears to be dominated by luminescence from the oxide, which is planned to be described in later publications. The additional emission resulting from roughness, indicated by the difference curve of Fig. 9, has a linear onset below a critical photon energy of about 4.5 eV, and falls off in intensity in the red, in a manner similar to that observed in light-emitting tunnel junctions. The fact that the emission spectrum changed dramatically with surface roughness indicates that the additional component is surface-plasmon mediated. The onset energy of 4.5
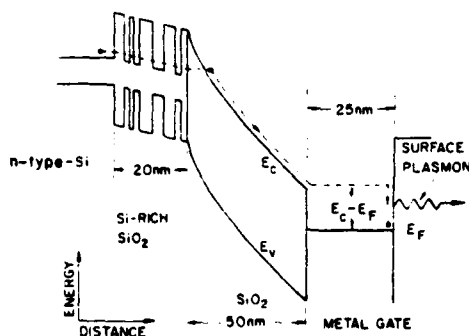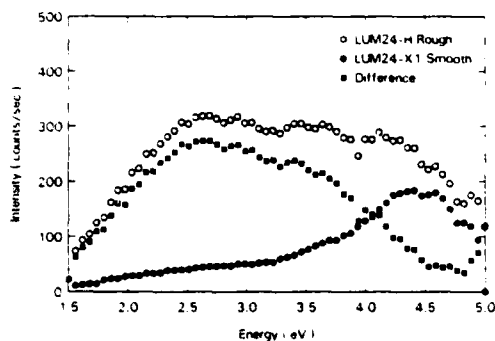
FIG. 9. Emission spectrum from a rough (LUM24-H,
open circles) vs a smooth (LUM24-X1, closed circles) Al-
gate charge-injection structure biased at 10 μA. The
difference spectrum, which is due to surface-plasmon-
coupled radiation, has a linear onset at 4.5 eV, corre-
sponding to the maximum energy, relative to the Fermi
surface of the metal, of the injected electrons.

FIG. 10. Emission spectrum from an Al-gate charge-
injection structure fabricated on a Si substrate with a grat-
ing profile. The sharp emission peak corresponds well to
the dip in the specular reflection caused by coupling to
Al-air interface surface plasmons. The total integrated ef-
ficiency of this peak emission is comparable to that ob-
served from light-emitting tunnel junctions on gratings,
when correction is made for the different dielectric prop-
erties of the two types of samples.

eV corresponds to the maximum energy (with
respect to the Fermi level) of electrons injected into
the metal electrode. This is larger than the potential
step of 3.6 eV between the bottom of the conduction
band of the SiO_2 and the Fermi level of the alumi-
num, obtained by Solomon and DiMaria[26] from
similar devices using internal photoemission. Thus
the spectral cutoff tells us that on the average. the
injected electrons are heated about 1 V above the
bottom of the oxide conduction band. This is the
first direct evidence for such heating, although it
has been predicted theoretically.[27]

The total external quantum efficiency for
LUM24-H was about 10^{-6} photons/electron, which
is comparable to randomly roughened tunnel junc-
tion efficiencies of 10^{-4}, when one takes into account
that the Al-gate electrode has much more highly
damped surface-plasmon polaritons than the Ag and
Au counter electrodes typically used for tunnel junc-
tions.

A more dramatic proof that part of the emission
from these charge-injection structures was surface-
plasmon mediated was obtained by fabricating them
on Si substrates with grating profiles. In Fig. 10 we
show the light emission from a charge-injection de-
vice with a 25-nm-thick Al gate, fabricated on an
815-nm periodicity, 36-nm amplitude holographical-
ly produced grating. The emission was measured
through a 7° wide and 16° high aperture. The spec-
trum shown in Fig. 10 was taken with an observa-
tion angle of 13° with respect to the sample normal.
The position in energy of the sharp emission peak
corresponded well to that predicted for the Al-
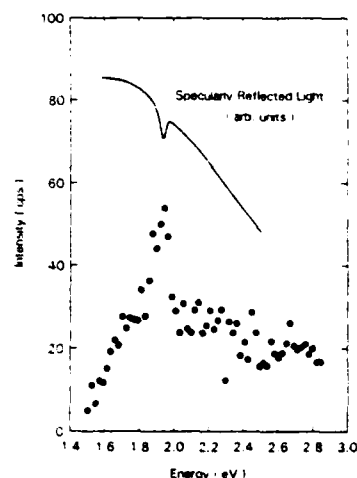vacuum interface fast surface-plasmon polariton.

Also included in Fig. 10 is a plot of the specularly
reflected radiation from a collimated beam from a
tungsten lamp directed at the Al grating of the
charge-injection structure, using the same geometry
and collection optics as for the emission experiment.
The reflectivity has a dip at the same energy as the
peak in the emission, due to absorption of the in-
cident radiation by surface plasmons. The specular-
ly reflected light is plotted on a logarithmic scale.
The dip corresponds to about a 90% decrease in the
reflected intensity in p polarization. The emission
peaks also varied in energy in the expected manner
as the observation angle was changed.

There is a continuum background in Fig. 10 that
was larger relative to the peak size than is typically
observed for emission from $Al-Al_2O_3-Ag$ tunnel
junctions on gratings. One reason was that the
surface-plasmon resonance is weaker at the Al-air
interface than at the Ag-air interface because of the
relative dielectric properties of the two metals. The
theory of Laks and Mills predicts a ratio of smooth
background to peak height of about 2 to 1 for the
Al-air interface plasmon emission at 2 V. We ob-
serve a ratio of about 1 to 1. There is probably also
a contribution to the residual background from cou-
pling caused by the random roughness of the

41

charge-injection layer, and also emission from the oxide itself.

The total integrated intensity under the sharp emission peak of Fig. 10 is 8.2×10^{-12} counts/electron sr. This is about 8 times smaller than would be predicted by analogy with our tunneling junction results using the theory of Laks and Mills[13] (see, for example, Fig. 10 of Ref. 6) to take into account the differences in dielectric properties of an Al-air interface versus a tunnel junction structure, grating amplitude, and grating periodicity. This difference could well result from the fact that the charge-injection structure inject at 3.5 eV, higher than the ~2.5 eV for the tunnel junctions studied by Kirtley et al. Electrons at higher energies have shorter lifetimes, resulting in lower external efficiencies. Also, the charge-injection structures were run at room temperature, while the tunnel junctions were usually run at low temperatures. Emission efficiencies are higher at lower temperatures, as we have shown above.

We interpret our results from the charge-injection structures as follows: Electrons in the conduction band of the SiO_2 have relatively short mean free paths (2—6 nm),[28] and are therefore primarily near (~1 eV) the bottom of the SiO_2 conduction band at the field strengths we are using.[27] They are injected into the Al with a narrow energy distribution, but relax before emitting surface plasmons, giving the same characteristic linear onset observed in light emission from tunnel junctions. The very similar emission efficiencies and spectral properties in the charge-injection structures and the tunnel junctions indicate that similar physical processes are occurring in each. Since the oxide layer is about 50-nm thick in the charge-injection structure, direct tunneling through the SiO_2 is certainly not a conduction mechansim in these devices, and inelastic tunneling cannot be the driving factor behind the emission.

V. CONCLUSIONS

The experimental evidence presented above argues for a hot-electron mechanism, as opposed to an in-elastic tunneling mechanism, for light emission from tunnel junctions, as well as from optically pumped thin films and charge-injection structures.

A very interesting puzzle arises from these results since light emission from tunnel junctions is sym-metric, or nearly so[1] with respect to the bias voltage polarity (we have confirmed this observation for emission from junctions on gratings), but surface-plasmon-mediated emission from charge-injection structures only occurs when the metal gate is held positive. These observations can be reconciled with the other experiments reported in this paper if the tunneling process creates hot electrons (or holes) on both sides of the insulating region. The convention-al one-electron picture of electron tunneling results in a hot electron being injected into one electrode, leaving a cold hole in the other. It may be that exci-tations on opposite sides of the insulating region are strongly coupled if the insulator is sufficiently thin, possibly through the slow surface-plasmon polariton mode.

Once suitable analysis methods have been developed, light emission from metal films will represent a probe of hot-electron dynamics in met-als, and charge transport through insulators, in an energy regime and with an energy resolution una-vailable to other techniques, and under a variety of different excitation conditions. Understanding the underlying mechanisms is therefore important for potential applications that extend well beyond the analysis of light emission from tunnel junctions. The hot-electron model is an aid in understanding these processes.

ACKNOWLEDGMENTS

[1] J. Lambe and S. L. McCarthy, Phys. Rev. Lett. 37, 923 (1976).

[2] S. L. McCarthy and J. Lambe, Appl. Phys. Lett. 30, 427 (1977).

[3] P. K. Hansma and H. P. Broida, Appl. Phys. Lett. 32, 545 (1978).

[4] A. Adams, J. C. Wyss, and P. K. Hansma, Phys. Rev. Lett. 42, 912 (1979).

[5] A. Adams and P. K. Hansma, Phys. Rev. B 23, 3597 (1981).

[6] J. R. Kirtley, T. N. Theis, and J. C. Tsang, Phys. Rev. B 24, 5650 (1981).

[7] J. R. Kirtley, T. N. Theis, and J. C. Tsang, Appl. Phys. Lett. 37, 435 (1980).

[8] D. Hone, B. Muhlschlegel, and D. J. Scalapino, Appl. Phys. Lett. 33, 203 (1978).

[9] R. W. Rendell, D. J. Scalapino, and B. Muhlschlegel, Phys. Rev. Lett. 41, 1746 (1978).

[10] L. C. Davis, Phys. Rev. B 16, 2482 (1977).

[11] B. Laks and D. L. Mills, Phys. Rev. B 20, 4962 (1979);

[11], 5175 (1980).

[12]E. N. Economou, Phys. Rev. 182, 539 (1969).

[13]B. Laks and D. L. Mills, Phys. Rev. B 22, 5723 (1980).

[14]L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968), p. 269.

[15]J. G. Simmons, J. Appl. Phys. 34, 1793 (1963).

[16]J. Moreland, A. Adams, and P. K. Hansma, Phys. Rev. B 25, 2297 (1982).

[17]C. R. Crowell and S. M. Sze, in *Physics of Thin Films,* edited by G. Hass and R. E. Thun (Academic, New York, 1967), Vol. IV, p. 325.

[18]H. J. Levinson, E. W. Plummer, and P. J. Feibelman, Phys. Rev. Lett. 43, 952 (1979).

[19]J. R. Kirtley, T. N. Theis, J. J. Chang, and D. L. Mills (unpublished).

[20]J. C. Tsang, J. R. Kirtley, and T. N. Theis, Solid State Commun. 35, 667 (1980).

[21]A. Girlando, W. Knoll, and M. R. Philpott, Solid State Commun. 38, 895 (1981).

[22]T. N. Theis, J. R. Kirtley, D. J. DiMaria, and D. W. Dong (unpublished).

[23]D. J. DiMaria, R. Ghez, and D. W. Dong, J. Appl. Phys. 51, 4830 (1980).

[24]D. J. DiMaria, K. M. DeMeyer, C. M. Serrano, and D. W. Dong, J. Appl. Phys. 52, 4825 (1981).

[25]A. Hartstein, J. C. Tsang, D. J. DiMaria, and D. W. Dong, Appl. Phys. Lett. 36, 836 (1980).

[26]P. M. Solomon and D. J. DiMaria, J. Appl. Phys. 52, 5867 (1981).

[27]H.-J Fitting and J.-U Friemann, Phys. Status Solidi 69, 349 (1982).

[28]R. Poirier and J. Oliver, Appl. Phys. Lett. 21, 334 (1972).

A study of the electrical and luminescence characteristics of a novel Si-based thin film electroluminescent device

D. J. Robbins,[a] D. J. DiMaria, C. Falcony,[b] and D. W. Dong

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

The results of electrical and luminescence measurements on a new, low voltage, dc, thin film electroluminescent device structure are presented. The devices incorporate a two-phase Si-rich-SiO_2/SiO_2 electron injector layer which provides control of current, and an active luminescent ZnS:Mn layer in which light is generated by hot electron impact excitation of the Mn^{2+} activator in high electric field. Separation of the processes of current control and light generation into different layers permits the effects of space charge and the average field distributions to be determined. The electroluminescence intensity is simply proportional to the average power dissipated in the ZnS:Mn layer when the average field is in the range 0.6–1.2 MV cm^{-1}, and when field distortion due to electron trapping in the SiO_2 layer is small. When the field locally in the ZnS:Mn layer exceeds ~ 1.7 MV cm^{-1}, lattice ionization competes with impact excitation of Mn^{2+} and the quantum efficiency falls. A simplified model assumes that in a quasi-steady-state condition the majority of the device current is carried by electrons accumulated in the satellite (L, X) valleys in the conduction band of ZnS. The results are compared with previous studies, and their general significance as regards the limiting efficiency of high field electroluminescent devices using ZnS is discussed.

PACS numbers: 78.60.Fi

I. INTRODUCTION

A recent letter from this laboratory[1] (hereafter referred to as Paper I) has described a novel dc thin film electroluminescent (DCTFEL) device which makes use of a two-phase Si-rich-SiO_2/SiO_2 electron injector layer as a solid state cathode. It was shown that the two-phase material, originally developed for use in electrically-alterable read-only-memory (EAROM) devices,[2] is able to deliver sufficient current density to make it useful in a solid state display device when used in conjunction with a manganese doped zinc sulphide (ZnS:Mn) luminescent layer. The light produced is the characteristic yellow $^4T_1 \rightarrow{}^6A_1$ intracenter d-d luminescence of the Mn^{2+} ion in ZnS.[3] The layer structure of a typical device is illustrated in Fig. 1(a), and an energy band diagram with the semitransparent gate electrode biased positively (i.e., forward biased) is shown in Fig. 1(b).

The most successful high field electroluminescence (EL) technology to date is that involving ac thin film electroluminescent (ACTFEL) structures,[4] using a ZnS:Mn layer sandwiched between two dielectric layers of high breakdown strength. Devices of this kind can have good brightness stability, but the capacitive coupling requires high voltage ac drive waveforms (typically > 100 V rms) which in turn necessitate high cost customized driving circuitry. From this point of view a lower voltage, DCEL technology has many attractions. However, it has long been recognized that one of the principal problems in producing a practical high field DCTFEL device lies in controlling the current to prevent

runaway and the resulting catastrophic breakdown.[5] Several methods of controlling current injection into the active luminescent layer have been tried; these include tunnel injection in reverse-biased metal-insulator-semiconductor (MIS) diodes[6] and in heterojunctions formed electrically[7] or during growth,[8,9] and the use of series resistive layers, both thin film[10] and powder.[11] None of the these approaches has so far produced a viable DCTFEL device technology, and this remains an active area of display research.

The performance and potential device advantages of the DCTFEL structures illustrated in Fig. 1(a) have been outlined in Paper I. Current control in these devices is achieved through a combination of electron tunneling processes in the two-phase injector layer (see Sec. III). The Si-rich-SiO_2 layer, which contains small (< 50 Å) Si islands in an SiO_2 matrix,[2] serves two main functions under forward bias. First, it screens against low-field breakdown due to imperfections at the Si substrate interface which forms one contact, and secondly it promotes electron tunneling into the SiO_2 layer at lower *average* fields (and hence lower voltage) than would be the case for a planar Si/SiO_2 interface. The stoichiometric SiO_2 layer is necessary to provide a barrier to electrons moving through the Si-rich-SiO_2 layer under forward bias. It is the field in the oxide near the interface with the Si-rich-SiO_2 material which controls electron injection from the last layer of Si islands into the SiO_2 via a Fowler–Nordheim tunneling mechanism,[2] and it is this tunneling process which provides control of current flow in the device as a whole. However, the use of a stoichiometric SiO_2 layer has one serious disadvantage which prevents practical application of these devices in their present form. The SiO_2 layer, although of good quality, contains deep electron traps in the bulk which gradually fill as electron current passes through

---
[a] Permanent address: Royal Signals and Radar Establishment, St. Andrews Road, Great Malvern, Worcs WR14 3 PS, Great Britain.
[b] Permanent address: Centre de Investigaciones del I.P.N., Dpto. de Fisica, Apdo. Post. 14-740, Mexico 14, D.F., Mexico.
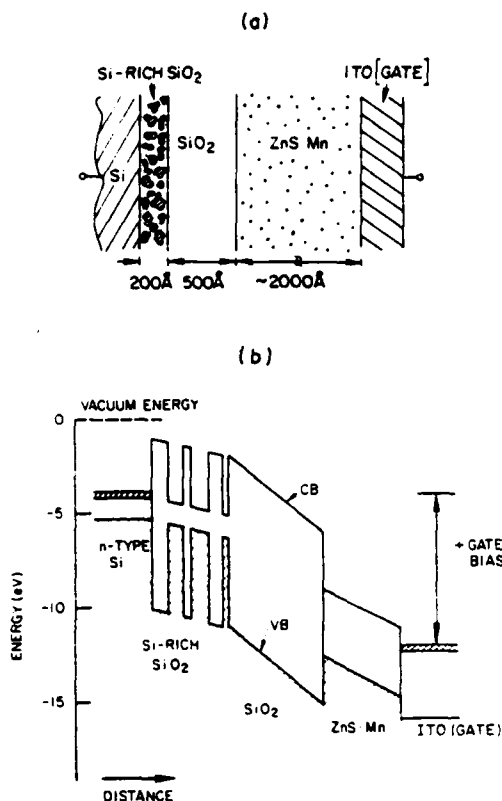
(a)

Si-RICH SiO₂ ... ITO[GATE]

(b)

FIG. 1. (a) Section through a typical DCTFEL device. (b) Energy band diagram under forward (positive gate) bias. The potential wells in the Si-rich-SiO₂ layer represent small (≲50 Å diameter) islands of Si. For simplicity, field distortion due to space-charge accumulation is not shown.

the device.[12] These traps are stable at room temperature, so that a permanent negative charge accumulates in the SiO₂ layer. This leads to current instability at constant forward bias voltage, and generates internal fields which result in catastrophic breakdown after a few hours dc operation.[1]

On the other hand, these devices represent a near-ideal structure for the study of the physical mechanisms underlying high field EL. A major advantage is the clear separation between the processes controlling current and those giving rise to the light emission. This allows electrons to be injected into the luminescent ZnS:Mn layer in a relatively well-defined manner, without need for carrier generation within the active layer by field ionization of traps or by avalanching processes, as is necessary in the ACTFEL devices.[13,14] In addition, the use of separate "control" devices to characterize the electron injector layer on each wafer, and the ability to detect changes in field at the Si-rich-SiO₂/SiO₂ interface via changes in device current, allow relationships between the space-charge distributions in the SiO₂ and ZnS:Mn layers to be established, and the average fields in the two layers to be determined. This detailed information about the electrical properties of the device provides useful insight into the factors influencing the luminescence characteristics of the ZnS:Mn layer.

Recently the possible advantages of "hot" electron in-

jection into the luminescent layer have been discussed for a structure in which the processes of electron heating and light generation are spatially separated.[15] One novel feature of the device illustrated in Fig. 1 is that in principal such "hot" electron injection is made possible by the potential energy step at the SiO₂/ZnS interface which arises from the difference in electron affinity between the two materials. Electrons injected from the conduction band (CB) of the SiO₂ under forward bias may increase the quantum efficiency of electroluminescence compared with electrons entering (or generated in) the ZnS:Mn layer near the CB minimum for ZnS.

The purpose of this paper is to present a detailed analysis of both the electrical and luminescence characteristics of these new DCTFEL devices. Section II gives experimental details concerning sample preparation and device measurement techniques. Section III summarizes the results of the various measurements which were carried out, including material analysis using secondary-ion mass spectroscopy (SIMS). An analysis of the electrical characteristics of the devices is presented in Sec. IV, and of the luminescence characteristics in Sec. V. Finally, the conclusions drawn from these analyses are discussed in Sec. VI. In particular, a simplified description of the high field EL process in ZnS:Mn is developed, based on concepts arising from earlier studies of MIS diodes, and is shown to give a consistent explanation of the major experimental observations.

II. EXPERIMENTAL DETAILS

A. Sample preparation

The electroluminescent devices, schematically illustrated in Fig. 1(a), were fabricated on ⟨100⟩, 0.001 Ω cm, n-type silicon substrates 1.25 in. in diameter. The Si-rich-SiO₂ and the SiO₂ layers were deposited by a chemical vapor deposition (CVD) process.[16] The excess silicon content in the Si-rich-SiO₂ was introduced by adjusting the ratio (R_0) of the concentrations of N_2O to SiH_4 in the gas phase. Deposition of stoichiometric SiO_2 requires $R_0 \approx 200$, while ≳ 13% excess atomic Si in the Si-rich oxide is obtained with $R_0 = 3$. The presence of water-related impurities was minimized by annealing the oxides at 1000 °C in N_2 for 30 min. The ZnS:Mn layer was deposited on the center portion of the wafer through a 1.5 × 1.5 cm mask using e-beam evaportion with separate sources for the ZnS and Mn components. The deposited layer was subsequently annealed at 600 °C in vacuum or in an N_2 ambient. The Mn:Zn mole ratio determined by x-ray fluorescence was typically ~0.005. Semitransparent circular indium tin oxide (ITO) electrodes approximately 5000 Å in thickness and 1.25 mm in diameter were then deposited in a 10 × 10 matrix array. The ITO was dc sputtered from an In_2O_3-SnO_2 target (91-9 M%) in an Ar ambient and annealed after deposition at 500 °C in an N_2 ambient. Some of the electrodes were deposited on the SiO₂ area not covered with ZnS:Mn, generating the control devices for electrical characterization of the injector layer. A set of devices (15N-4) was fabricated without the Si-rich-SiO₂ injector layer present in order to measure the capacitance versus voltage (C-V) characteristics of these devices.
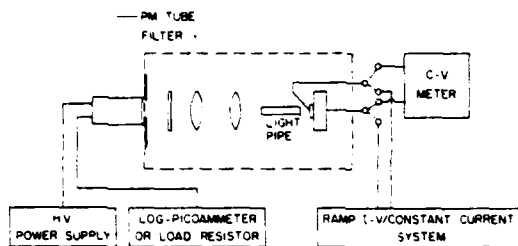
4554   J. Appl. Phys., Vol. 54, No. 8, August 1983

45

Robbins et al.   455

FIG. 2. Schematic diagram of the experimental apparatus.

B. Measurement techniques

The experimental arrangement for the electrical and optical measurements reported in this work is illustrated in Fig. 2. The sample was positioned on a vertical vacuum chuck and a spring-loaded thin wire probe was used to make electrical contact on the top electrode (ITO gate). The emitted light was collected with a 3-mm-diam quartz pipe and a set of two convex lenses. An RCA 1P21 photomultiplier tube was used to detect the emitted light. This light was passed through a yellow cutoff filter. The sample and the optics were enclosed in a light-tight box flushed with N_2.

A ramped current as a function of voltage (ramp I-V) technique and a constant current injection technique have been used to investigate the electrical and luminescence behavior of the EL devices. The ramp I-V characteristics were obtained using a constant voltage-ramp-rate magnitude of 0.5 V sec^{-1} applied to the sample. The electron current flowing through the device and the light emission intensity were simultaneously monitored as a function of the applied voltage with two log-picoammeters (Keithley model No. 26000). The relative voltage displacements of the ramp I-V characteristics were used for calculations of the amount of trapped charge in the oxide and/or the ZnS:Mn layers. The constant current injection technique uses a feedback network that automatically adjusts the applied voltage to compensate for any changes of a preset average current. Two modes of this technique were used, one with a dc applied voltage and the other with a triangular voltage waveform as the applied bias. Both the incremental voltage required to maintain constant average current and the photomultiplier output were continuously monitored. Capacitance versus voltage $(C$-$V)$ measurements were performed on devices without Si-rich-SiO$_2$ layers to ensure a good substrate Si-SiO$_2$ interface, with the minimum interface state density required for this type of measurement. A Boonton capacitance meter (model number 72-BD) and a ramp voltage generator were used for these measurements.

III. RESULTS

A. Ramp I-V measurements

Current versus voltage $(I$-$V)$ curves, measured with a voltage source ramped at a constant rate of 0.5 V sec^{-1} from 0 V, for different gate electrode locations on sample 9L-14 are shown in Fig. 3(a). The arrows indicate directions of increasing and decreasing ramp voltage. The constant-current
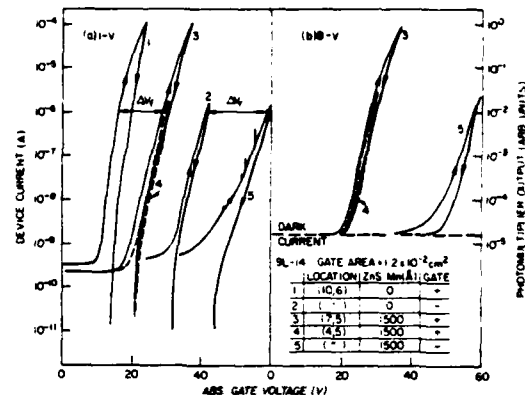

FIG. 3. (a) Current-voltage $(I$-$V)$ characteristics under forward (+ gate) and reverse (− gate) bias for wafer 9L-14. Gate electrode polarity and location in the 10 × 10 gate matrix are given in the inset. (b) Light output-voltage $(B$-$V)$ curves measured simultaneously with the I-V curves in (a) bearing the corresponding reference number.

portion of the curves represents a displacement current I_d arising from the capacitance C of the structure; i.e., $I_d = CdV/dt$. When the voltage ramp is reversed to decrease the applied voltage, the measured current eventually falls to a very low value and ultimately changes sign as I_d changes sign with the reverse ramp.

Curves 1 and 2 refer to control devices on the wafer, outside the area with the ZnS:Mn layer. Curve 1 is measured under forward (+ gate) bias, and curve 2 under reverse (− gate) bias. The hysteresis evident in the I-V characteristics arises from electron trapping in the SiO$_2$ layer (see Sec. IV). Forward bias (curves 3,4) and reverse bias (curve 5) I-V characteristics for devices including the ZnS:Mn layer are also illustrated in Fig. 3(a). It is shown in Sec. IV that, to a good approximation, the voltage drop across the ZnS:Mn layer at any current and for a particular gate polarity is given by the voltage difference at that current between the control device and the device including the ZnS:Mn layer. Hence ΔV_f and ΔV_r in Fig. 3(a) represent the voltages across the ZnS:Mn layer under forward and reverse bias respectively, at a current of 10^{-6} A.

The devices behave as electrical diodes due to the properties of the two-phase Si-rich-SiO$_2$/SiO$_2$ injector layer. Real particle current (i.e., $I > I_d$) is controlled by tunneling at a limiting interfacial energy barrier for either gate polarity. Under forward bias, current is limited by the ~3.1 eV barrier at the Si/SiO$_2$ interface arising from the electron affinity difference between Si (4.0 eV) and SiO$_2$ (0.9 eV). Under reverse bias, the current-limiting barriers occur at the ZnS:Mn/SiO$_2$ interface (~3.0 eV), or, for the control devices, at the ITO/SiO$_2$ interface (~2.8 eV). This latter value represents an effective barrier height on these devices, and was obtained by fitting the Fowler–Nordheim tunneling equation to the reverse bias I-V characteristic of a control device. This barrier height is lower than that found[17] for In$_2$O$_3$ on planar SiO$_2$/Si (~4 eV), probably because the granular texture introduced by the underlying Si-rich-SiO$_2$ layer produces some small field enhancement at the injecting

ITO/SiO$_2$ interface, thereby reducing the effective barrier by ~1 eV. Current injection by the Fowler–Nordheim mechanism at a planar, homogeneous interface with effective barrier height ~3 eV requires a field of $\gtrsim 5$ MV cm^{-1} in the barrier layer.[18] For an SiO$_2$ thickness of 500 Å, as used in these devices, this field corresponds to $\gtrsim 25$ V across the oxide layer as threshold for real particle current injection. This is evident in curve 2 for the control device under reverse bias. Under forward bias, however, injection occurs not at a planar Si/SiO$_2$ interface, but from the last layers of small Si islands in the Si-rich-SiO$_2$ layer. Field enhancement at surface irregularities on those islands[19] produces injection at an *average* field of only ~2 MV cm^{-1} in the SiO$_2$ layer, corresponding to the threshold $(I > I_d)$ of ~10 V in curve 1 of Fig. 3(a).

In addition to allowing real current injection at low forward bias, and thereby generating the diode character of these devices, the Si-rich-SiO$_2$ layer is valuable in screening against low-field breakdown.[1] Under reverse bias it is not possible to pass high currents $(\sim 10^{-4}$ A) without rapid breakdown; curve 5 in Fig. 3(a) shows current spikes due to localized breakdown which are quite typical for reverse bias operation. Under forward bias however the space charge accumulating in the small Si islands of the Si-rich-SiO$_2$ layer effectively screens against local breakdown caused by irregularities at the interface with the n-type Si substrate. This produces macroscopically-uniform current injection, and allows sustained high current operation by preventing local current runaway.

Figure 3(b) shows light-output versus voltage (*B-V*) curves measured simultaneously with the *I-V* curves bearing the same reference number in Fig. 3(a). Luminance measured using a calibrated spot photometer gives an average external luminous efficiency of ~0.3 lm/W^{-1} for these devices.[1] This efficiency compares favorably with other EL technologies, and has been achieved without systematic optimization of the devices. Because of the macroscopically-uniform current injection provided by the Si-rich-SiO$_2$ layer, the luminescence from the ZnS:Mn layer is generally similarly uniform; any nonuniformity tends to arise from imperfections in the deposition processes for the ZnS:Mn and/or ITO layers.

B. Constant current measurements

Constant current measurements (dc) were made by applying a ramp voltage source (0.5 V sec^{-1}) to produce a current of 1×10^{-4} A under forward bias, and then switching in the feedback circuitry to maintain that current by increasing the applied voltage. The results for wafer 9L-14 are summarized in Fig. 4. The left-hand ordinate records the incremental voltage ΔV_g required to maintain $I_g = 1 \times 10^{-4}$ A as a function of time, plotted for both a control device and a device with ZnS:Mn layer. The right-hand ordinate records the relative light output for the luminescent device at constant current, measured simultaneously with ΔV_g(ZnS).

As current passes through these devices, electrons are trapped in the SiO$_2$ layer producing a space charge which reduces the field near the injecting Si-rich-SiO$_2$/SiO$_2$ inter-
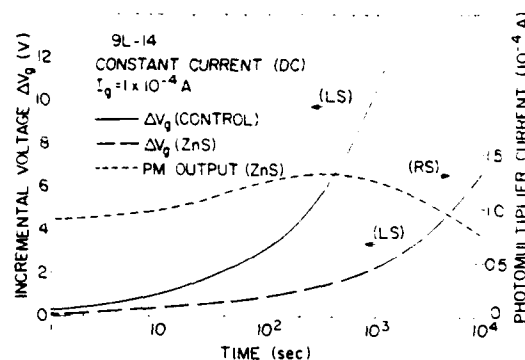

FIG. 4. Time-dependent measurements on wafer 9L-14 at a constant current $I_g = 1 \times 10^{-4}$ A. Changes in electric field distribution resulting from space-charge accumulation cause the light output to vary, although the device current is held constant.

face.[12] The incremental voltage ΔV_g is applied to overcome the effects of this space charge and maintain a constant field near the injecting interface. Because of the electron trapping the device has no steady-state condition even with constant current drive, as is evident from the form of the light-output curve in Fig. 4. After some period of time, which is shorter the higher the current, the field generated by the trapped charge becomes sufficiently large to cause catastrophic breakdown of the device; this occurred after ~20 min operation at $I_g = 1 \times 10^{-4}$ A for control devices on wafer 9L-14, when ΔV_g(control) ~12 V. It is clear from Fig. 4 that at any given time, ΔV_g(ZnS) $< \Delta V_g$(control), i.e., the rate at which the applied voltage must be increased to maintain constant current is lower for the devices with a ZnS:Mn layer than for control devices. This is due to the accumulation of positive space charge in the ZnS:Mn layer which partially screens the negative space charge trapped in the SiO$_2$ layer. The luminescent devices can be run for a longer period of time (~5 h) before catastrophic breakdown occurs, probably due to the increase in dielectric thickness introduced by the ZnS:Mn layer. However breakdown tends to occur when ΔV_g(ZnS) ~10 V, similar to the value for the control devices.

C. *C-V* measurements

An independent estimate of the space charge accumulation in the ZnS:Mn layer has been made from *C-V* measurements on sample 15N-4, shown in Fig. 5. This wafer has no Si-rich-SiO$_2$ layer, and the gate electrodes are thick Ni rather than ITO dots. The Si/SiO$_2$/ZnS/Ni structure represents a MIS device with a dual dielectric layer,and the capacitance of the planar Si/SiO$_2$ interface can be used as a probe of charge trapping in the composite dielectric layer. The *I-V* characteristics of wafer 15N-4 are essentially similar to those for 15N-5 previously reported in Paper I, the latter having ITO gate electrodes. The use of thick Ni is preferred for the *C-V* measurements since it avoids any contact problems associated with possible depletion of the ITO interface region; however measurement of *C-V* curves for 15N-5 showed behavior essentially similar to 15N-4, so the results are not very sensitive to the gate electrode material.

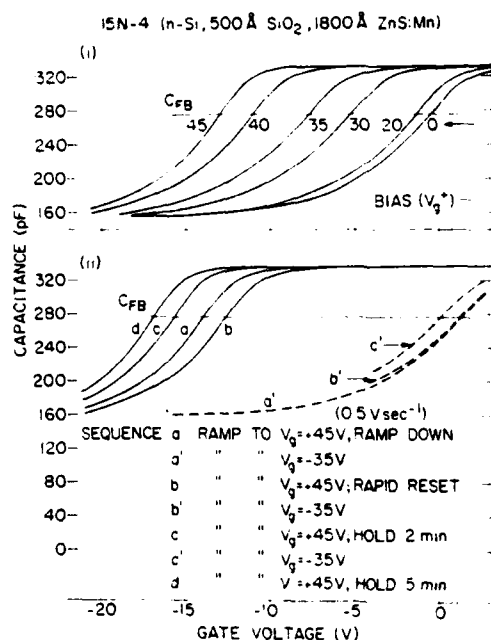4556   J. Appl. Phys., Vol. 54, No. 8, August 1983

47

Robbins *et al*   4556

FIG. 5. Capacitance-voltage (C-V) measurements on wafer 15N-4, with the indicated layer structure. The gate electrodes were thick Ni. The structure is that of an MIS device with a dual dielectric (SiO$_2$/ZnS:Mn) layer. (a) C-V curves measured after a positive gate bias (V_g^-) of the indicated magnitude had been applied by a voltage ramp, and then removed. (b) The effect of applying a positive gate bias V_g^- = 45 V, followed by a negative bias V_g^- = − 35 V. The sequence of curves is indicated in the figure.

The dual SiO$_2$ (500 Å)/ZnS:Mn (1800 Å) layer is equivalent to a single SiO$_2$ dielectric ~ 1400 Å thick, deposited on 2 Ω cm, n-type Si with a donor concentration ~2.5 × 10^{15} cm^{-3}. For an ideal MIS capacitor with this structure, measured at high frequency,

$$C_{FB}/C_i \sim 0.84, \tag{1}$$

$$C_{min}/C_i \sim 0.48, \tag{2}$$

where C_{FB} is the flat-band capacitance, and, for an n-type semiconductor, C_i and C_{min} are the upper and lower capacitance limits with positive and negative gate bias, respectively.[20] For wafer 15N-4, from Fig. 5(a), we have:

$$C_i = 330 \text{ pF}, \tag{3}$$

$$C_{min} = 155 \text{ pF}. \tag{4}$$

Therefore, C_{min}/C_i = 0.47, close to the ideal MIS value. Using Eq. (1) the flat-band capacitance for 15N-4 is estimated to be C_{FB} = 277 pF, and this value is indicated in Fig. 5(a).

The series of curves in Fig. 5(a) were obtained in sequence, beginning with V_g^- = 0, by first applying a positive voltage ramp to the gate electrode, sweeping up to the voltage indicated and sweeping down to zero, before finally taking the C-V curve. For an n-type semiconductor the C-V curves shift in a direction consistent with positive space-charge accumulation in the dielectric layer, the amount of charge increasing with increasing V_g^-. Control devices out-

side the ZnS:Mn layer do not show this effect, demonstrating that the positive space charge is associated with the luminescent layer. I-V measurements[1] on the electrically-equivalent wafer 15N-5 show that for V_g^- <45 V there is little real particle current (< 10^{-8} A); electron trapping in the SiO$_2$ layer is therefore insignificant in these C-V measurements. The C-V curves are analyzed quantitatively in Sec. IV.

Figure 5(b) shows the results of a sequence of measurements in which a positive gate voltage (V_g^- = 45 V) is first applied for a given period of time, a C-V curve taken, then a negative gate voltage (V_g^- = − 35 V) is applied, and finally a second C-V curve taken. The sequence is given in detail in the bottom part of Fig. 5(b). The accumulation of positive charge in the ZnS:Mn layer increases when V_g^- is applied for longer times. The negative bias V_g^- = − 35 V is insufficient to produce real particle current flow through the entire device, but it does return C_{FB} to approximately 0 V. The negative bias therefore effectively annihilates the accumulated positive space charge in the ZnS:Mn layer, presumably via recombination involving electrons injected at the Ni electrode.

D. Pulsed addressing

Further insight into the diode properties of the devices and into the mechanism of light emission can be obtained by applying voltage pulses rather than dc bias to the gate electrodes. Figure 6 illustrates the gate current and light-output waveforms for single, 5-msec unipolar pulses of both polarities. Under forward bias, with V_{gp}^+ = 31 V [Fig. 6(a)], the
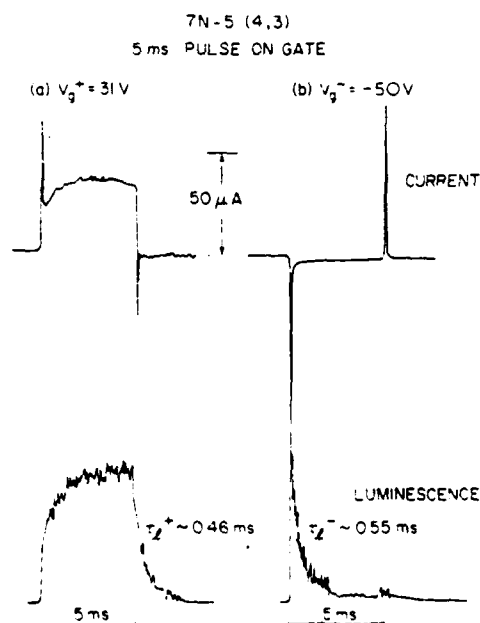

FIG. 6. Typical current (upper) and light-output (lower) curves on applying a single 5-msec unipolar pulse to the gate electrode. (a) Forward bias. (b) Reverse bias. The time taken for the luminescence to decay to 1/e of its peak value after the end of the pulse is also indicated.

48

gate current shows a displacement current spike followed by a real particle current which rises towards a steady state with a slow (~1 msec) time constant. The shape of this gate current waveform is sensitive to the pulse voltage and length, to the number of unipolar pulses applied, and can be changed drastically if negative bias pulses are applied between the positive pulses V_{gp}^{-+}.[21] The corresponding light output curve rises towards a steady state value, following the gate current, and decays after the end of the pulse with a time constant τ_l^{-+} ~0.46 msec.

Figure 6(b) shows similar measurements for reverse bias pulses with $V_{gp}^{-} = -50$ V, the higher reverse voltage being necessary to generate detectable light output. The current waveform is reversed in polarity, and is dominated by the displacement current spikes; there is negligible real particle current. The luminescence waveform is quite different from that in the forward direction. It shows a peak corresponding to the displacement spike on switching on the pulse, but then decays through the duration of the pulse with a time constant τ_l^{-} ~0.55 msec. There is a second small increase in light output at the end of the pulse, arising from relaxation of the polarized charge distribution as the voltage pulse is switched off. The decay of the luminescence waveform τ_l^{+} and τ_l^{-} is determined largely by the intrinsic Mn^{2+} luminescence decay time, giving τ_{Mn} ~0.5 msec for the ZnS:Mn material used in these devices.

The C-V measurements have demonstrated that positive gate bias generates positive space-charge accumulation in the ZnS:Mn layer, and that this space charge can be removed by application of negative gate bias. In order to investigate the effects of this positive space charge on the efficiency of light output, a comparison has been made of addressing by unipolar and by bipolar pulses under conditions of constant average gate current. Triangular voltage waveforms were used for this purpose, as illustrated in the insets to Figs. 7(a) and 7(b). Figure 7(a) plots the voltage peak V_p (left ordinate) and the corresponding photomultiplier output (right ordinate) as a function of time when the voltage is pulsed positive from 0 V and the average gate current is maintained at $\langle I_g \rangle = 1 \times 10^{-5}$ A. The voltage control circuit is switched on at time $t = 0$, increases to a value V_p ~43 V to produce the required current, and then increases slowly with time to overcome electron trapping in the oxide and maintain that current. Figure 7(b) shows similar plots, but with a bipolar voltage waveform pulsed positive from a negative bias level of -35 V. The voltage peak required to maintain the gate current $\langle I_g \rangle = 1 \times 10^{-5}$ A is now V_p ~80 V, giving a peak value ~45 V above 0 V. The fact that the positive part of the pulse waveform is similar to the unipolar case [Fig. 7(a)] confirms that the current is dominated by the forward bias, and that negligible real particle current flows in reverse bias under these conditions. However the application of the negative bias increases the light intensity by ~2 for the same average gate current. This arises from a reduction in the positive space charge accumulating in the ZnS:Mn layer in the presence of the negative bias, which increases the average field, and therefore the useful power, in this luminescent layer.
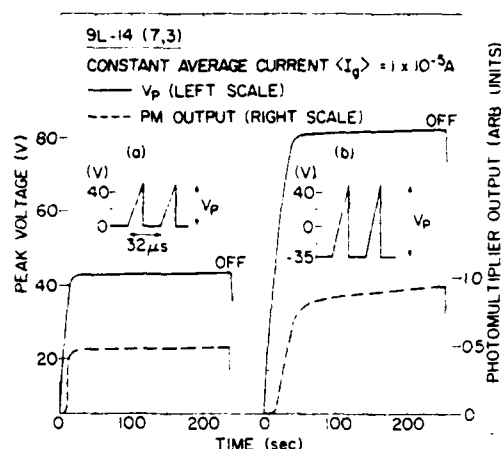

FIG. 7. Time-dependent measurements on an active device on wafer 9L-14 at a constant *average* current $\langle I_g \rangle = 1 \times 10^{-5}$ A, using the triangular voltage waveforms indicated in the insets and a feedback circuit. (a) Voltage pulsed positive from 0 V. (b) Voltage pulsed positive from a negative bias level -35 V.

E. Composite structure and SIMS analysis

In order to determine whether EL efficiency is influenced by hot electron injection into the ZnS:Mn layer over the potential energy step at the SiO_2/ZnS interface, comparative measurements have been made on the devices illustrated at the top of Fig. 8. Device 9L-11 was made with
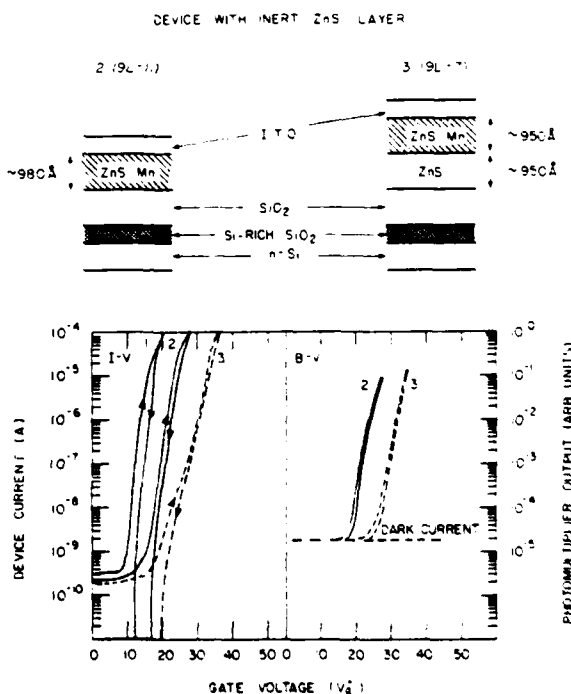
FIG. 8. (Upper) The layer structure of two wafers used to investigate the effect of electron injection at the SiO_2/ZnS interface. Wafer 9L-7 incorporates an inert layer of ZnS without Mn^{2+} activator adjacent to the SiO_2 layer. (Lower) I-V and B-V measurements on these two wafers. Curve 1 is the I-V characteristic of control devices on both wafers. Curves 2 and 3 refer to active devices on wafers 9L-11 and 9L-7, respectively.

49

the normal structure, except that the ZnS:Mn layer was only ~980 Å thick, i.e., about half the normal thickness. Device 9L-7 had an additional layer of ~950 Å of ZnS *without* Mn deposited on the SiO₂ before deposition of a further 950 Å of luminescent ZnS:Mn material. Both wafers were annealed at 500 °C for 1 h under N₂ before deposition of the gate electrode matrix. Therefore, as prepared, the light-emitting layers on the two wafers were similar. The Mn concentration in both was determined by x-ray fluorescence to be ~1%, a factor of ~2 higher than in the thicker standard ZnS:Mn layers. This higher Mn concentration is expected to reduce the luminous efficiency.

If hot electron injection at the SiO₂/ZnS interface takes place it should in principle affect the luminescence efficiency only for wafer 9L-11, since for 9L-7 the ZnS layer adjacent to the SiO₂ was made nonluminescent. Before conclusions can be drawn, however, it is necessary to determine the extent of Mn redistribution during post-growth annealing. A SIMS analysis of both wafers has therefore been made in the area between the gate electrodes to determine the Mn profile[22]; the results are shown in Fig. 9. The steep increase in the Si count rate defines the position of the ZnS/SiO₂ interface in these plots.

A number of conclusions can be drawn from these results. First, it is clear that in both wafers the Mn distribution is not uniform, and that there is evidence of Mn accumulation at the surface. Secondly, Mn has diffused into the undoped ZnS layer on wafer 9L-7. This sample shows three distinct regions up to the SiO₂ interface: (i) the region of surface Mn accumulation, (ii) a region of uniform Mn concentration corresponding to the doped ZnS:Mn layer, (iii) a

region of steadily decreasing Mn concentration up to the SiO₂ interface, corresponding to the nominally-undoped ZnS layer into which Mn has been introduced by diffusion from the outer doped layer. For wafer 9L-11 the count ratio of Mn/Zn is >10 at the SiO₂ interface; for wafer 9L-7 the same count ratio of Mn/Zn is <1. It can therefore be concluded that for wafer 9L-7 the Mn concentration near the interface is at least one order of magnitude lower than for 9L-11. Finally, the apparent tailing of the Mn and Zn profiles into the SiO₂ layer may indicate interdiffusion which could be significant with respect to device performance, particularly life. This tailing is, for example, much greater than observed for ZnS:Mn layers deposited directly on Si substrates.[23] However it is possible that the effect is the result of recoil during the sputtering process, combined with a higher ionization probability for the Mn and Zn atoms in the polar SiO₂ matrix when compared with elemental Si. This point is undergoing further investigation.

The bottom half of Fig. 8 shows I-V and B-V curves for devices on the two wafers. The voltage required to produce a current of 10^{-4} A is higher for 9L-7 (curve B) because of the greater overall ZnS thickness on this wafer, but it is clear that the light output at any given current is very similar for both wafers. This is discussed further in Sec. VI.

IV. ELECTRICAL CHARACTERISTICS

A. Field distribution

The electroluminescent devices investigated here represent a dual-dielectric structure sandwiched between the Si-rich-SiO₂ layer and the gate electrode. Following Ref. 12, the electrical characteristics can be analyzed in terms of the elec-
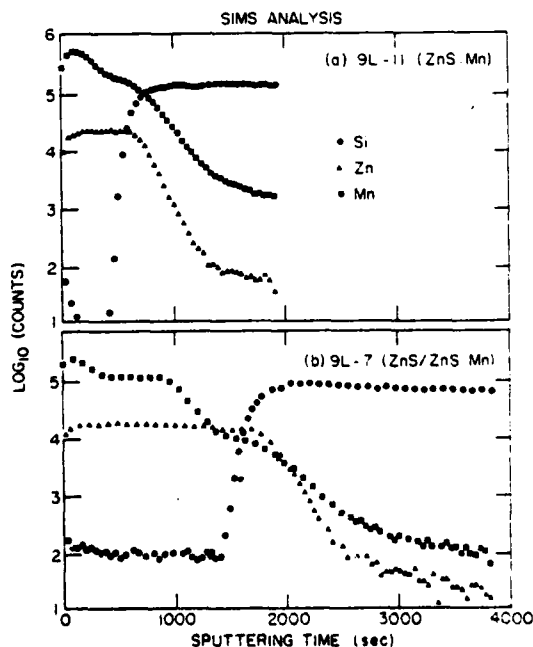
FIG. 9. Comparison of SIMS analyses of the two structures illustrated in Fig. 8 (upper). (a) Wafer 9L-11. (b) Wafer 9L-7. Sputtering by Ar⁺ ions was in an area between the ITO gate electrodes.
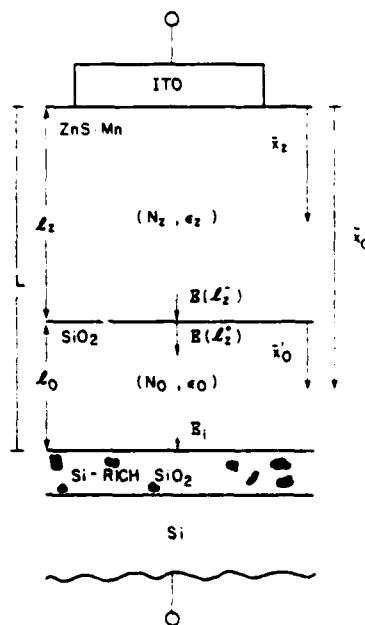
FIG. 10. Schematic diagram of electric field (E) and areal space-charge density (N) distributions in the devices illustrated in Fig. 1(a), under forward bias. The quantities $\bar{x}_z$ and $\bar{x}_0$ are the centroids of the positive and negative space-charge distributions in the ZnS:Mn and SiO₂ layers, respectively. All distances refer to the gate electrode/ZnS:Mn interface as origin.

4559   J. Appl. Phys., Vol. 54 No. 8, August 1983

Robbins *et al.*   4559

50

tric field and space-charge distributions illustrated schematically in Fig. 10. From Poisson's equation, the electrostatic field at any position x in the two layers is given by

$$E_z(x) = E(0) + \int_0^x \frac{\rho(x')dx'}{\epsilon_z} \quad (0 < x < l_z) \tag{5}$$

for the ZnS:Mn layer, and

$$E_0(x) = \frac{\epsilon_z}{\epsilon_0} E(0) + \int_{l_z}^x \frac{\rho(x')dx'}{\epsilon_0} \quad (l_z < x < L) \tag{6}$$

for the SiO$_2$ layer, where $E(0)$ is the electrostatic field in the ZnS:Mn layer at the interface with the gate electrode $(x = 0)$. The subscripts 0 and z refer to the SiO$_2$ and ZnS:Mn layers respectively, and $\rho(x)$ is the space-charge density. After integration one finds

$$E(0) = \frac{(V_g - \phi_{ms} - \Psi_s) - Q_0\left(\frac{l_z + l_0 - \bar{x}_0}{\epsilon_0}\right) - Q_z\left(\frac{(l_z - \bar{x}_z)}{\epsilon_z} + \frac{l_0}{\epsilon_0}\right)}{l_z + (\epsilon_z/\epsilon_0)l_0}, \tag{7}$$

where V_g is the gate voltage, ϕ_{ms} is the work function difference between the gate electrode material and Si substrate, Ψ_s is the Si surface potential, $Q_0(Q_z)$ is the space charge per unit area in the SiO$_2$ (ZnS:Mn) layer, $\bar{x}_0(\bar{x}_z)$ is the centroid of the charge distribution in the SiO$_2$ (ZnS:Mn) layer measured from the gate electrode interface, and $\epsilon_0(\epsilon_z)$ is the low frequency permittivity of SiO$_2$ (ZnS:Mn). For the samples used in this work, ϕ_{ms} and Ψ_s are small compared with V_g and will generally be neglected. Also, the voltage dropped across the Si-rich-SiO$_2$ layer is negligible compared to V_g for the electric field conditions of interest in this study.[12] Other useful field relationships are:

(1) For the field at each side of the SiO$_2$/ZnS:Mn interface

$$E_0(l_z^+) = \frac{\epsilon_z}{\epsilon_0} E(0) + \frac{Q_z}{\epsilon_0}$$
$$= \frac{\epsilon_z}{\epsilon_0} E_z(l_z^-); \tag{8}$$

(2) For the field at the Si-rich-SiO$_2$/SiO$_2$ interface

$$E(L) = \frac{\epsilon_z}{\epsilon_0} E(0) + \frac{Q_0 + Q_z}{\epsilon_0}, \tag{9}$$

where $L = (l_0 + l_z)$ is the total dielectric thickness.

The field E_i in the SiO$_2$ close to the Si-rich-SiO$_2$ interface controls current injection into the dielectric layers under forward bias. It is known that the real particle current is carried by electrons, the current density being determined by the Fowler–Nordheim tunneling relationship[12,18]:

$$J = aE_i^2 \exp[-b/E_i], \tag{10}$$

where E_i can be taken as approximately equal to $E(L)$. Figure 11 shows data from a typical ramp I-V measurement on a control device on wafer 9L-14 plotted according to Eq. (10), with $E_i = (V_g^+/l_0)$. Agreement is good except at the highest fields, where the current density becomes sufficiently high that significant electron trapping occurs in the SiO$_2$ layer. Although the interfacial field is assumed to be equivalent to the average electric field (for the case of no charge trapping for a planar contact to SiO$_2$), the actual interfacial field is larger by approximately a factor of 2. This field enhancement is believed to be caused by the size, shape, and density of Si islands in the Si-rich-SiO$_2$ layer near the SiO$_2$/Si-rich-SiO$_2$ interface.[12] However, this field enhancement factor can be treated, to first order, as a constant and absorbed into the

Fowler–Nordheim constants a and b.[12] This treatment for the field enhancement will be followed throughout this article.

Using the superscripts c and z to denote devices without and with the ZnS:Mn active layer respectively, the average fields in the SiO$_2$ and ZnS:Mn layers are related to the applied gate voltages by

$$\bar{E}_0^c l_0 = V_g^c - \phi_{ms} - \Psi_s, \tag{11}$$

for the control devices, and

$$\bar{E}_z^z l_z + \bar{E}_0^z l_0 = V_g^z - \phi_{ms} - \Psi_s, \tag{12}$$

for the luminenscent devices. Using simple electrostatics, these average fields can be expressed as follows in terms of the current-limiting field at the Si-rich-SiO$_2$/SiO$_2$ interface and the space-charge distribution:

$$\bar{E}_0 = \int_{l_z}^L E(x)dx/l_0$$
$$= E(L) - \frac{Q_0\bar{x}_0'}{\epsilon_0 l_0}$$
$$= \frac{(V_g - \phi_{ms} - \Psi_s) + \frac{Q_0 l_z}{\epsilon_z l_0}(l_0 - \bar{x}_0') + \frac{Q_z\bar{x}_z}{\epsilon_z}}{l_0 + \frac{\epsilon_0}{\epsilon_z}l_z} \tag{13}$$
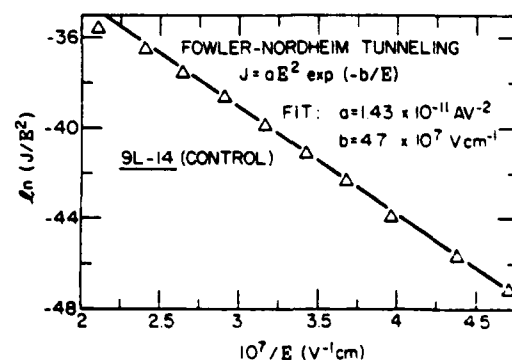
FIG. 11. A plot of the I-V characteristic of a control device on wafer 9L-14 [e.g., curve 1 in Fig. 3(a)] according to the Fowler–Nordheim tunneling equation. The electric field E is taken as the average oxide field $\bar{E}_n$. The deviation between the extrapolated line and the experimental data at high current density is used as a measure of electron trapping in the SiO$_2$ layer.

4560   J. Appl. Phys., Vol. 54, No. 8, August 1983

Robbins et al.   4560

51

and

$$\bar{E}_z = \int_0^{l_z} E(x)dx/l_z$$

$$= \frac{\epsilon_0}{\epsilon_z}E(L) - \frac{Q_0}{\epsilon_z} - \frac{Q_z\bar{x}_z}{\epsilon_z l_z}$$

$$= \frac{(V_g - \phi_{ms} - \Psi_s) - \frac{Q_0}{\epsilon_0}(l_0 - \bar{x}_0') - \frac{Q_z\bar{x}_z l_0}{\epsilon_0 l_z}}{l_z + \frac{\epsilon_z}{\epsilon_0}l_0} , \quad (14)$$

where $\bar{x}_0' = (\bar{x}_0 - l_z)$. For measurements on control and luminescent devices at identical currents, $E^c(L) = E^z(L)$. In addition, since electron trapping in the oxide layer should be largely independent of the ZnS:Mn layer, it will be assumed that the space charge accumulating in the SiO_2 is proportional only to the integrated charge passed through the devices, i.e., for equivalent voltage ramps and at equal currents, $Q_0^c = Q_0^z$. Hence, at a given current and for the same oxide charge, Eq. (13) gives

$$\bar{E}_0^z = \bar{E}_0^c \quad (15)$$

for luminescent and control devices on the same wafer. Substituting in Eqs. (11) and (12), and neglecting small terms in ϕ_{ms} and Ψ_s, gives the following expressions for the average fields in terms of the measured gate voltages:

$$\bar{E}_0 = V_g^c/l_0, \quad (16)$$

$$\bar{E}_z = (V_g^z - V_g^c)/l_z, \quad (17)$$

where V_g^c and V_g^z are measured at equal currents and for equal integrated charge passed through devices on the same wafer. The circles in Fig. 12, referred to the left-hand ordinate, represent a plot of the average field in the ZnS:Mn layer $(\bar{E}_z)$ versus the average field in the SiO_2 layer $(\bar{E}_0)$ for an
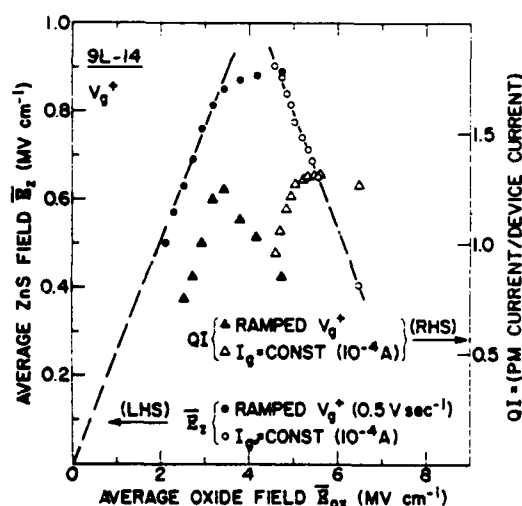

FIG. 12. Average field in the ZnS:Mn layer $(\bar{E}_z$, left-hand ordinate) and the relative luminescence quantum intensity (QI, right-hand ordinate) as a function of average field in the SiO_2 layer. The data points are derived from the analysis of both ramp I-V and constant current measurements on wafer 9L-14 under forward bias, as described in the text.

active device under forward bias, calculated according to Eqs. (16) and (17) and using data from both ramp I-V and constant current measurements on wafer 9L-14.

The result [Eq. (17)] that the voltage drop across the ZnS:Mn layer is given by the voltage difference $(V_g^z - V_g^c)$, with V_g^z and V_g^c measured at the same current, was derived for forward bias conditions. In that case, the current for both control and active devices is determined by the same interface field, approximately $E(L)$. In reverse bias, however, the current in the control devices is determined by tunneling at the ITO/SiO_2 interface, and in the active devices by tunneling at the ZnS/SiO_2 interface. It was pointed out in Sec. III A that the effective barrier heights at these two interfaces are in practice very similar. Essentially parallel arguments can then be used to show that Eq. (15) is also applicable in reverse bias, and therefore that to a good approximation the average field in the ZnS:Mn layer under reverse bias is also given by Eq. (17).

B. Space-charge distribution

Electron trapping in the SiO_2 layer at a fixed gate bias V_g^+ reduces the interface field $E(L)$, and hence causes the current to decrease. An estimate of charge trapping in the oxide layer during a ramp I-V measurement can be made from the Fowler–Nordheim plot in Fig. 11. When trapping is negligible (i.e., $Q_0 \sim 0$), $E(L) = \bar{E}_0$ from Eq. (13). A plot of $\ln[J/(\bar{E}_0)^2]$ vs $(\bar{E}_0)^{-1}$ should therefore be linear at low current density, and show deviations from linearity when Q_0 becomes significant: in Fig. 11 this deviation occurs for device currents $> 10^{-6}$ A. The experimental points in Fig. 11 are plotted according to the average field $\bar{E}_0$, whereas the extrapolated linear fit defines a dependence on the interface field, approximately $E(L)$. Neglecting the small contribution from the pre-exponential factor in the Fowler–Nordheim tunneling equation, the difference δE between the experimental data and the linear fit at the same current density J is given approximately by

$$\delta E|_J \approx \bar{E}_0 - E(L)$$

$$\approx -\frac{Q_0\bar{x}_0'}{\epsilon_0 l_0}. \quad (18)$$

For electron trapping $Q_0 = -qN_0$, where q is the electronic charge and N_0 the areal space-charge density in the SiO_2 layer. From independent measurements on oxide layers of the kind used here it is known[2] that the space charge is uniformly distributed in the SiO_2, i.e., $\bar{x}_0' = l_0/2$. Therefore, from the ramp I-V measurements on control devices, the trapped electron charge in the SiO_2 layer can be determined at a particular current density from the expression

$$\delta E|_J \approx \frac{qN_0}{2\epsilon_0}. \quad (19)$$

In constant current measurements ($I_c = 10^{-4}$ A in Fig. 4), the gate voltage V_g^+ is increased to offset the negative space charge in the oxide and maintain the field $E(L)$ constant, i.e., after a time t

52

$$\Delta E(L) = E(L)|_t - E(L)|_{t=0}$$

$$= \Delta \bar{E}_0 + \frac{\Delta Q_0 \bar{x}_0'}{\epsilon_0 l_0}$$

$$= 0 \text{ at constant current.} \qquad (20)$$

Substituting for $\bar{E}_0$, Q_0 and $\bar{x}_0'$ gives

$$\Delta V_g^c = \frac{q \Delta N_0 l_0}{2\epsilon_0}, \qquad (21)$$

i.e., the change in areal space-charge density ΔN_0 can be calculated directly from the incremental voltage ΔV_g^c required to maintain a constant current in the control devices. The total oxide charge under these conditions is obtained by adding to ΔN_0 the amount of charge accumulated during initial ramping of V_g^- to reach the required gate current, $I_g = 10^{-4}$ A. Equations (19) and (21) have been used to calculate the values of N_0 plotted in Fig. 13 as a function of average oxide field $\bar{E}_0$, using both ramp I-V and constant current measurements on wafer 9L-14.

When N_0 is known, the space charge in the ZnS:Mn layer can be calculated by combining Eqs. (13)–(17). Since the C-V measurements described in Sec. III C demonstrate that the ZnS space charge is positive, we have

$$Q_z = qN_z, \qquad (22)$$

and for the ramp I-V measurements

$$\bar{E}_z = \frac{V_g^z - V_g^c}{l_z}$$

$$\approx \frac{V_g^z + \frac{qN_0 l_0}{2\epsilon_0} - \frac{qN_z \bar{x}_z l_0}{\epsilon_0 l_z}}{l_z + \frac{\epsilon_z}{\epsilon_0} l_0}, \qquad (23)$$

where the terms ϕ_{ms} and Ψ_s are again neglected. Hence

$$\frac{N_z \bar{x}_z}{l_z} = \frac{\epsilon_z}{q l_z}(V_g^c - V_g^z) + \frac{\epsilon_0 V_g^c}{q l_0} + \frac{N_0}{2}. \qquad (24)$$

For the constant current measurements, $\Delta E(L) = 0$. Hence from Eq. (14)

$$\frac{\Delta N_z \bar{x}_z}{l_z} = -\frac{\epsilon_z}{q} \Delta \bar{E}_z + \Delta N_0$$

$$= \frac{\epsilon_z}{q l_z}(\Delta V_g^c - \Delta V_g^z) + \Delta N_0. \qquad (25)$$

Equations (24) and (25) have been used to calculate the values of $N_z \bar{x}_z / l_z$ plotted in Fig. 13 as a function of $\bar{E}_0$. For the ZnS:Mn layer there is no available independent information which can be used to fix the ratio $\bar{x}_z / l_z$, so the experimentally-determined quantity is the centroid-weighted space-charge density $N_z \bar{x}_z / l_z$.

Inspection of the experimental data in Figs. 12 and 13 shows that there are certain simple relationships between the space-charge distributions and the average fields. These are:

(1) Low current (ramp I-V data):

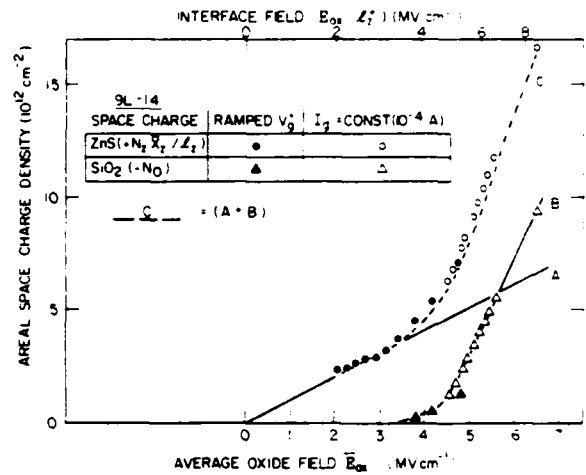$$\bar{E}_z \sim \frac{\bar{E}_0}{4}, \qquad (26a)$$

FIG. 13. Areal space-charge density accumulated in both the ZnS:Mn and the SiO_2 layers as a function of average field in the SiO_2. The data points are derived from the analysis of both ramp I-V and constant current measurements on wafer 9L-14 under forward bias, as described in the text. The negative space charge in the bulk of the SiO_2 layer is known to be uniformly distributed. The distribution of positive space charge in the ZnS:Mn layer is not known, so that only the centroid-weighted areal density $N_z \bar{x}_z / l_z$ can be determined from the data. The top abscissa scale shows the field in the oxide at the SiO_2/ZnS:Mn interface, which is higher than the average field when electron trapping in the SiO_2 is significant.

$$\frac{N_z \bar{x}_z}{l_z} \sim \frac{\epsilon_z \bar{E}_0}{4q}. \qquad (26b)$$

(ii) High current (constant current, $I_g = 10^{-4}$ A):

$$\Delta \bar{E}_z \sim -\frac{\Delta \bar{E}_0}{4}, \qquad (26c)$$

$$\frac{\Delta N_z \bar{x}_z}{l_z} \sim \frac{\epsilon_z \Delta \bar{E}_0}{4q} + \Delta N_0. \qquad (26d)$$

The following equality is obtained by simultaneously solving Eqs. (13) and (14) for the interface field $E(L)$

$$\frac{\epsilon_z \bar{E}_z}{q} + \frac{N_z \bar{x}_z}{l_z} = \frac{\epsilon_0 \bar{E}_0}{q} + N_0 \frac{(l_0 - \bar{x}_0')}{l_0}. \qquad (27)$$

When one of the pair of equations [Eqs. (26a) and (26b)] is specified, the other follows from this equality by virtue of the fact that $N_0 \sim 0$ at low current and $\epsilon_0 \sim \epsilon_z/2$. Similarly, when one equation of the pair Eqs. (26c) and (26d) is specified the other follows from Eq. (27) by using Eq. (20), and making the substitution $\bar{x}_0' = l_0/2$. Hence, the form of the expressions (26b) and (26d) for the space charge in the ZnS:Mn is sufficient to determine the average field relationships [Eqs. (26a) and (26c)], or vice versa.

The form of the experimental data can be rationalized in the following way. For low currents, when $N_0 \sim 0$, the space charge in the ZnS:Mn layer follows line A in Fig. 13. This line is given by Eq. (26b), which can be rewritten as

$$\frac{N_z \bar{x}_z}{l_z} \sim \frac{\epsilon_0 \bar{E}_0}{2q}, \qquad (28)$$

since $\epsilon_z \sim 8$ and $\epsilon_0 \sim 3.9$. If the positive space charge were

4562   J. Appl. Phys., Vol. 54, No. 8, August 1983

Robbins et al.   4562

53

uniformly distributed throughout the ZnS:Mn layer (as is the case for the negative space charge in the oxide), then this would reduce to

$$N_z \sim \frac{\epsilon_0 \overline{E}_0}{q} \text{(with } \bar{x}_z = l_{z/2} \text{ assumed).} \tag{29}$$

Therefore line A in Fig. 13 would be predicted for a uniform space-charge distribution in the ZnS:Mn layer, with an areal density N_z equal to the average electric displacement in the SiO_2 layer. Under conditions of constant current ($I_g = 10^{-4}$ A), Eq. (21) gives an expression for the change in the oxide trapped charge:

$$\Delta N_0 = \frac{2 \Delta V_g^c \epsilon_0}{q l_0}. \tag{30}$$

A similar expression for the centroid-weighted space charge in the ZnS layer is obtained by substitution of Eq. (30) in Eq. (25), using thickness parameters appropriate to wafer 9L-14:

$$\frac{\Delta N_z \bar{x}_z}{l_z} = \frac{(3.93 \Delta V_g^c - \Delta V_g^z) \epsilon_z}{q l_z}. \tag{31}$$

Since it is evident in Fig. 4 that $3.93 \Delta V_g^c > \Delta V_g^z$ at equal times in this constant current measurement, Eq. (31) reduces to

$$\frac{\Delta N_z \bar{x}_z}{l_z} \approx 1.3 \Delta N_0. \tag{32}$$

Using Eq. (20) for constant current conditions gives

$$\frac{\epsilon_0 \Delta \overline{E}_0}{2q} = \frac{\Delta N_0}{4}. \tag{33}$$

Therefore combining Eqs. (32) and (33) for constant current gives the approximate relationship

$$\frac{\Delta N_z \bar{x}_z}{l_z} \approx \frac{\epsilon_0 \Delta \overline{E}_0}{2q} + 1.05 \Delta N_0. \tag{34}$$

Hence values of the weighted space-charge density $N_z \bar{x}_z / l_z$ calculated from the constant current data in Fig. 4 are predicted to follow quite closely curve $C = A + B$ when plotted in Fig. 13, where $A \equiv \epsilon_0 \overline{E}_0 / 2q$ and $B \equiv N_0$. This prediction is verified experimentally. Equations (28) and (34) are roughly equivalent to Eqs. (26b) and (26d), respectively, and are sufficient to determine the relationship between the average fields $\overline{E}_0$ and $\overline{E}_z$ represented by equations (26a) and (26c) and demonstrated experimentally in Fig. 12.

The magnitude of $N_z \bar{x}_z / l_z$, the centroid-weighted space-charge density in the ZnS:Mn layer deduced from the ramp I-V and constant current measurements, can be independently checked from the C-V data in Fig. 5. The flat-band shift for wafer 15N-4 is given by[12]

$$\Delta V_{FB} = q N_0 \left(\frac{\bar{x}_0'}{\epsilon_0} + \frac{l_z}{\epsilon_z} \right) - \frac{q N_z \bar{x}_z}{\epsilon_z}. \tag{35}$$

For this sample, with no Si-rich-SiO_2 layer, negligible current is passed through the devices for $V_g^+ < 45$ V. Hence $N_0 \sim 0$, and

$$\Delta V_{FB} \sim - \frac{q N_z \bar{x}_z}{\epsilon_z}. \tag{36}$$

For $V_g^+ = 45$ V, $\Delta V_{FB} = -13$ V after rapid reset of the ramp voltage, and $\Delta V_{FB} = -18$ V after V_g^+ has been held at 45 V for 5 min. These flat-band shifts correspond to weighted space-charge densities $N_z \bar{x}_z / l_z = 3.2$ and 4.4×10^{12} cm^{-2}, respectively. For $\Delta V_{FB} = -18$ V and $V_g^+ = 45$ V, the average field in the oxide layer is $\overline{E}_0 \sim 4.5$ MV cm^{-1}. From the data in Fig. 13 for wafer 9L-14, the quantity $N_z \bar{x}_z / l_z$ predicted from line A at this field is $\sim 4.8 \times 10^{12}$ cm^{-2}, and the total weighted density from line C is $\sim 6.1 \times 10^{12}$ cm^{-2}. Therefore agreement between the C-V measurements on wafer 15N-4 and the I-V measurements on wafer 9L-14 is quite reasonable as regards the magnitude of positive space charge. Also it should be borne in mind that some difference between the two techniques might be expected from the fact that negligible current flows through the devices on wafer 15N-4 during the voltage ramping, whereas a current $I_g \sim 10^{-4}$ A flows through the devices on wafer 9L-14 when $\overline{E}_0 \sim 4.5$ MV cm^{-1}.

V. LUMINESCENCE CHARACTERISTICS

Figure 14 is a logarithmic plot of photomultiplier current versus device (gate) current for various gate locations on two separate wafers; the data for 9L-14 are taken from the ramp I-V curves in Fig. 3. For wafer 9L-14 the slope of this plot is close to unity, showing that the luminescence efficiency per unit current (i.e., the quantum efficiency, QE) does not vary greatly over the current and field range represented, except for a small decrease at the highest currents. Comparing the data points for gate location (4,5) on wafer 9L-14 under positive and negative gate bias shows that the QE is not significantly different for the two bias directions. This was generally found to be the case for those wafers producing the highest light output. In some cases, e.g., wafer 7N-5, the QE was observed to be higher with forward bias than with reverse, but with lower overall efficiency compared with the best wafers. The QE measured in this way is sensitive to electron trapping in the devices, since trapping affects the field distribution. However the fact that the best

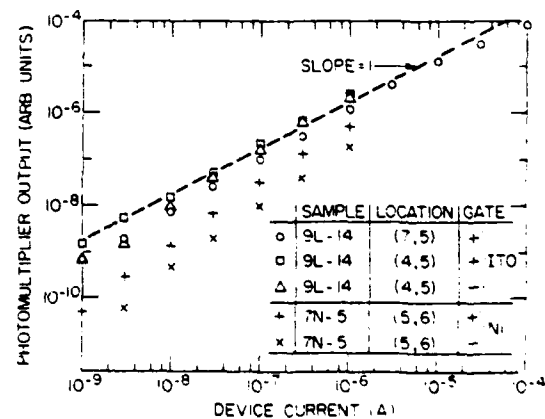| SAMPLE	LOCATION	GATE
o 9L-14	(7,5)	+
□ 9L-14	(4,5)	+ ITO
△ 9L-14	(4,5)	-
+ 7N-5	(5,6)	+ Ni
x 7N-5	(5,6)	-

FIG. 14. A logarithmic plot of relative light output vs device current for the samples indicated in the inset, taken from ramp I-V data.

54

devices, which had minimal trapping rates, showed equivalent QE under both forward and reverse bias suggests that electron injection over the potential step at the $SiO_2/$ZnS:Mn interface under forward bias has little effect on the luminescence efficiency.

More insight into the luminescence mechanism can be obtained by plotting the light intensity per unit current (i.e., the quantum intensity, QI) as a function of average oxide field $\overline{E}_0$. This is done in Fig. 12 for both ramp I-V and constant current data from wafer 9L-14, referred to the right-hand ordinate. At low currents and fields, the QI increases in parallel with the average field in the luminescent ZnS:Mn layer $(\overline{E}_z)$. Therefore, in the range $\overline{E}_z \sim 0.6$–0.9 MV cm^{-1}, the QI is a linear function of $\overline{E}_z$ and the luminescence intensity is directly proportional to the average power dissipated in the ZnS:Mn layer. At a field $\overline{E}_0 \sim 3.4$ MV cm^{-1} the current becomes sufficiently large that significant electron trapping in the oxide occurs. At this point the ZnS:Mn field $\overline{E}_z$ ceases to increase in proportion to $\overline{E}_0$; it is initially roughly constant, and then falls sharply as the rate of positive space charge accumulation in the ZnS:Mn layer increases. At the same time the onset of electron trapping in the SiO_2 layer triggers a sharp drop in the quantum intensity at high currents, as recorded in the ramp I-V measurements. This is the cause of the apparent saturation of light output for the high current points in Fig. 14. However this drop in QI can be reversed if the high current is sustained, as in the constant current measurements with $I_g = 10^{-4}$ A (Fig. 4). The QI under constant current plotted in Fig. 12 recovers towards its higher value even though the average field in the ZnS:Mn layer is decreasing because of the buildup of positive space charge. Eventually the QI reaches a maximum and begins to fall again as $\overline{E}_z$ continues to decrease. The effect of these changes is seen in the dotted curve in Fig. 4, where it is clear that the charge trapping in the oxide layer prevents attainment of a steady state in these devices. It is also clear that the luminescence efficiency is sensitive to the detailed field distribution in the ZnS:Mn layer, and not simply to the average field $\overline{E}_z$.

VI. DISCUSSION

A. Power dependence of luminescence

It was shown in the previous section that in the field regime $\overline{E}_z = 0.6$–0.9 MV cm^{-1} under forward bias the light output from the ZnS:Mn layer is simply proportional to the average power dissipated in that layer. The lower limit in these experiments is set by the detectability of luminescence against the dark current of the photomultiplier tube; the upper limit under forward bias is determined by the onset of electron trapping in the SiO_2 which generates significant field distortion in the ZnS:Mn layer and triggers a fall in QI. However, the measurements under reverse bias show that the proportionality between light output and field extends to average fields as high as $\overline{E}_z \sim 1.2$ MV cm^{-1}. For wafer 9L-14, using the data in Fig. 3, the voltage drop across the active ZnS:Mn layer at a current $|I_g| = 10^{-6}$ A is $\Delta V_f = 13$ V in forward bias, and $\Delta V_r = 18$ V in reverse bias; the average fields in forward and reverse bias are, therefore, $|\overline{E}_z| = 0.87$ and 1.2 MV cm^{-1}, respectively. At the same absolute cur-

rent $|I_g| = 10^{-6}$ A, the photomultiplier output is $B_f = 1.4 \times 10^{-2}$ in forward bias and $B_r = 2 \times 10^{-2}$ in reverse bias, in arbitrary units. Hence, considering both bias directions, the light output is closely proportional to the average power dissipated in the ZnS:Mn layer at least in the range $0.6 < |\overline{E}_z| < 1.2$ MV cm^{-1}, when the field distribuion is not severely distorted by charge trapping in the oxide layer.

A similar proportionality between light output and average power in the active layer is also evident in the data for the devices with composite ZnS/ZnS:Mn layers, illustrated in Fig. 8. For wafer 9L-11, with ~ 980 Å of ZnS:Mn, the voltage drop across the active layer in forward bias is ~ 7.8 V at $I_g = 10^{-4}$ A, giving an average field $\overline{E}_z \sim 0.80$ MV cm^{-1} at this gate current. Wafer 9L-7 has a total ZnS thickness of ~ 1900 Å, deposited in two steps as ~ 950 Å of ZnS followed by ~ 950 Å or ZnS:Mn. In forward bias the voltage drop across this whole ZnS layer is ~ 15.5 V at $I_g = 10^{-4}$ A, giving an average field $\overline{E}_z \sim 0.82$ MV cm^{-1}. The results of the SIMS analysis (Fig. 9) show that after annealing wafer 9L-7, most of the Mn dopant remains in the second ZnS layer next to the gate electrode, and that this second layer has a composition very similar to the active layer of wafer 9L-11. Some diffusion of Mn into the first ZnS layer on wafer 9L-7 does occur during annealing, but near the ZnS/SiO_2 interface the Mn concentration is more than an order of magnitude lower than in the second ZnS:Mn layer. Comparing light output for the two wafers at the gate current $I_g = 10^{-4}$ A gives, for 9L-11, $B_f = 1 \times 10^{-1}$ with $\overline{E}_z \sim 0.8$ MV cm^{-1}, and for 9L-7, $B_f = 1.1 \times 10^{-1}$ with $\overline{E}_z \sim 0.82$ MV cm^{-1}, where the average field in the latter case refers to the total ZnS/ZnS:Mn composite layer. Because of the lower luminescence efficiency of these layers the measurements refer to relatively high gate currents, for which some electron trapping in the oxide will occur. The field distribution in the ZnS layers may, therefore, be distorted by such trapping. However, the most probable explanation of the close correspondence between the two wafers is that the majority of the luminescence from 9L-7 is produced in the second ZnS:Mn layer, which is similar in thickness and composition to that on 9L-11, and that the average field in the second layer is approximately equal to that in the total ZnS thickness. The implication is then that similar light output results from both 9L-7 and 9L-11 when the same average power is dissipated in the doped ZnS:Mn layer. In particular, it does not appear important whether this active doped layer is adjacent to the SiO_2 layer, or displaced from it by a region of ZnS with very low Mn concentration.

B. The SiO_2/ZnS:Mn interface

It was pointed out in Sec. II that the electroluminescent devices studied here incorporate a potential energy step at the SiO_2/ZnS:Mn interface which in principle allows forward bias injection of electrons into the active layer at an energy ~ 3 eV above the CB minimum of ZnS. The effect of this hot electron injection might be observable as an increase in QE of luminescence in forward bias compared with reverse bias on the same device, or when comparing two c-

vices under forward bias where one incorporates the potential step for the active layer and the other does not. The discussion in Sec. VI A has shown that no such differences in QE have been observed, and that the light output in the field range $0.6 < \overline{E}_z < 1.2$ MV cm^{-1} is determined largely by the average power dissipated in the active layer, regardless of the bias direction or the method of electron injection into that layer.

However, an increase in excitation efficiency can apparently be realized in other systems when electrons are injected over a band edge step. Electrons injected from SiO$_2$ into Si have an excess kinetic energy of ~ 3.2 eV due to the difference in electron affinities between the two materials. Measurements of current in a Si p-n junction following electron injection from an SiO$_2$ layer suggest that each electron after entering the Si generates on average one electron-hole pair,[24] i.e., the excess kinetic energy can be used to ionize the Si lattice. Similarly, calculations show that electron ionization rates in avalanche diodes based on GaAs/Ga$_{1-x}$Al$_x$As multilayered heterojunction structures can be enhanced by electron injection over the CB step between Ga$_{1-x}$Al$_x$As and GaAs.[25] However, in both these materials, SiO$_2$/Si and Ga$_{1-x}$Al$_x$As/GaAs, the heterojunction interface is known to be of good quality. On the other hand, the first-to-grow layer of ZnS electron-beam evaporated onto an oxide dielectric (BaTiO$_3$) has been shown to have a very fine-grain structure with poor luminescence properties[26]; annealing at temperatures up to 550 °C, similar to those used here, produced no significant change in microstructure.[26,27] In our present device one reason for the absence of any enhancement in QE due to injection at the potential step between SiO$_2$ and ZnS:Mn is, therefore, likely to be the poor quality of the luminescent material near the heterojunction interface. Considerable improvement in deposition techniques or in annealing procedure may be necessary before such "hot" electron injection effects can be observed. In addition, however, a relatively efficient process of electric-field-induced electron heating in ZnS would restrict any observable enhancement of QE due to hot electron injection to very thin ZnS layers, and/or to low average fields. This point is discussed further in Sec. VI D.

C. Comparison with previous work

It has been shown in Paper I that the external luminous efficiency of these devices (~ 0.1–0.4 lm W^{-1}) is comparable with the best achieved with other EL technologies, both powder and thin film. The dependence of light output on current described here (Fig. 14) is similar to that observed in a number of thin film and epitaxial EL systems, where the relationship $B \propto I^n$ has been reported[8,9,28–30]; n is found to take a value close to, or just greater than, unity at low current, and to decrease at higher currents. More insight into the physics of the luminescence process is obtained by considering the quantum intensity QI ($= B/I$) as a function of field. Our results under forward bias, illustrated in Fig. 12 and discussed in Secs. V and VI A, show that $B/I \propto \overline{E}_z$ for average fields $\overline{E}_z > 0.6$ MV cm^{-1} and for low current conditions in which electron trapping in the SiO$_2$ layer is negligi-

ble. At higher currents the oxide trapping triggers a change in the distribution of space charge which tends to "pin" the average field $\overline{E}_z$, and which causes the ratio B/I to reach a maximum and subsequently decrease despite the continuing increase in bias V_g. A similar behavior is observed at high current in reverse-biased MIS diodes based on n-type ZnS:Mn (Refs. 31,32) and ZnSe:Mn (Ref. 33) crystals, the quantum efficiency reaching a maximum and then decreasing with increasing bias voltage. However, the low current behavior of these MIS diodes is very different, showing a strong *exponential increase* in the ratio B/I with applied bias voltage, in contrast to the linear dependence at low voltage evident in Fig. 12.

The marked difference in low current behavior between these MIS diodes and the forward-biased EL structures described in this work is somewhat surprising, since both involve electron injection by tunneling through an insulating layer into a high field region in insulating or depleted ZnS:Mn (or ZnSe:Mn). One reason for the discrepancy might be that the field in the active layer of the MIS diodes is significantly lower than that in our devices, and closer to the threshold for luminescence; Gordon[32] shows the exponential variation of B/I to occur for average fields in the depletion layer in the range 0.3–0.5 MV cm^{-1} for ZnS:Mn. This would also be consistent with the lower efficiency of the MIS diodes.[31]

D. The mechanism of light emission

It is now generally believed that electroluminescence in ZnS:Mn is generated primarily by impact excitation of the Mn^{2+} centers by hot electrons, rather than, for example, by transfer of the energy of recombination of electron-hole pairs created by avalanching.[34–36] The fact that in the devices studied here the light output per unit power dissipated in the ZnS:Mn layer has reached its peak value at a uniform average field as low as 0.6 MV cm^{-1} favors impact excitation as the dominant mechanism, since this would be a very low field for avalanche processes to dominate. Indeed, it is suggested below that the decrease in quantum intensity observed at higher applied voltages (Fig. 12) results from the onset of lattice ionization processes, which would then appear to be in competition with, rather than contributing to, the Mn^{2+} luminescence.

The explanation of the experimental results developed below rests firmly on the concept that the conduction band structure of ZnS (and ZnSe) provides an energy "trap" (region 2 in Fig. 15) in which electrons heated in an electric field rapidly accumulate, and from which it is difficult to escape because of strong intervalley scattering. The potential importance of the satellite CB valleys in hot electron phenomena in ZnSe is made clear by the observation of Gunn-type current oscillations,[37] with a threshold field of 3.8×10^4 V cm^{-1}. Later Livingstone and Allen[38] invoked higher-valley effects to explain the donor-density dependence of phonon scattering length derived from measurements of ionization coefficients as a function of field. Recently Gordon and Allen[39] proposed that the approximate resonance between the Mn^{2+} excitation energy and the $\Gamma - X_1$ energy

4565   J. Appl. Phys. Vol. 54 No. 8 August 1983

56

Robbins *et al*   4565
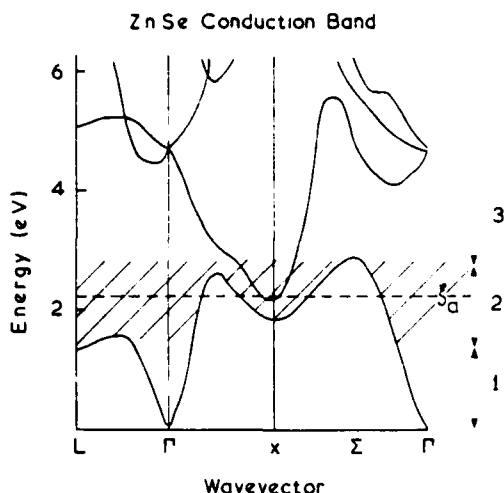
ZnSe Conduction Band

FIG. 15. Conduction band structure of ZnSe taken from Ref. 40, shown partitioned into the energy regions 1–3 discussed in the text. ZnS is expected to be similar. ξ_a represents the threshold for impact excitation of the Mn^{2+} activator in ZnSe or ZnS. The shading indicates the region of the band structure (region 2) in which electrons heated in an electric field will tend to accumulate because of strong intervalley scattering.

separation in the CB of ZnSe and ZnS leads to a strong Auger quenching of the Mn^{2+} luminescence in material with appreciable concentration of free carriers. This Auger quenching is the reverse process to impact excitation:

$$Mn^{2+}(^6A_1) + e_{CB}(X) \rightleftarrows Mn^{2+}(^4T_1) + e_{CB}(\Gamma). \quad (37)$$
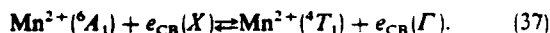
Subsequently, Gordon[32] pointed out that efficient electroluminescence in ZnSe occurs in a field regime approximately an order of magnitude larger than the Gunn effect threshold, where efficient transfer of electrons to the satellite valleys is expected. In a review of work on ZnSe:Mn MIS diodes, Allen[34] has suggested that the drop in quantum efficiency observed at high fields in those devices results from heating of electrons above the X_3 CB minimum, with consequent loss of the resonance implied by Eq. (37). These ideas form a basis for explaining the experimental results presented here.

A simple theory of impact excitation in MIS diodes has been given by Allen[33,34] in terms of an impact cross-section σ_E, which is a phenomenological parameter derived from an averaging over the electron energy distribution. The impact excitation rate $R(\xi)$ as a function of electron energy is given by:

$$R(\xi) = \sigma(\xi) v(\xi) N_a, \quad (38)$$

where $v(\xi)$ is the velocity of an electron with energy ξ, and N_a is the activator (Mn^{2+}) concentration. The microscopic cross-section $\sigma(\xi)$ is generally taken to be a step function with some threshold energy ξ_a for the onset of excitation. The experimental observables are convolutions of the one-electron quantities with the electron distribution function $f(\xi,E)$. The field-dependent cross-section σ_E is then defined by

$$\sigma_E(E)\langle v \rangle = \langle \sigma v \rangle$$

$$= \int_0^\infty \sigma(\xi) v(\xi) f(\xi,E) d\xi. \quad (39)$$

Under steady-state conditions, equating excitation and relaxation rates between the Mn^{2+} excited and ground states, and for nonsaturating current densities, Allen finds for the reverse-biased diodes[34]

$$\frac{B}{IW} \propto \sigma_E N_a \eta_r, \quad (40)$$

where W is the diode depletion width, and η_r the radiative efficiency for the activator excited state which is assumed independent of field. As noted in Sec. VI C, a plot of the quantity B/IW vs V^{-1} shows an exponential dependence at low voltage, which is attributed to an exponential increase in σ_E:

$$\sigma_E = \sigma_0 \exp(-E_s^2/E^2). \quad (41)$$

The value of E_s has been given as 8×10^5 V cm^{-1} for ZnS:Mn, and 6×10^5 V cm^{-1} for ZnSe:Mn diodes.[31] For a reverse-biased Schottky diode the depletion width W and the maximum field in the depletion region, E_{max}, are related by[20]

$$W \propto E_{max} \propto V^{1/2}. \quad (42)$$

Equation (40) can, therefore, be rewritten as

$$\frac{B}{IE_{max}} \propto \sigma_E N_a, \quad (43)$$

where η_r is assumed constant. However, the discussion in Sec. VI A has shown that the analogous relationship derived from our data takes the form

$$\frac{B}{IE_z} \sim \text{constant} \quad (0.6 < \overline{E}_z < 1.2 \text{ MV cm}^{-1}). \quad (44)$$

Since the cross-section analysis leading to Eq. (43) is quite generally applicable, the implication of the experimental data is that in the quoted field range, $0.6 < \overline{E}_z < 1.2$ MV cm^{-1}, the cross-section σ_E is approximately constant in our devices, i.e., the fraction of the electron distribution lying above the impact excitation threshold ξ_a is roughly constant in this field range. Tornquist and Tuomi[30] have reached a similar conclusion, although they attribute the constancy of the hot electron fraction above threshold to a "clamping" of the field. This is clearly not the case in our devices.

The observed constancy of the cross section combined with a variation in the field can be rationalized most easily by an extension of the band structure arguments originally proposed by Allen and co-workers. Figure 15 illustrates the calculated conduction band structure for ZnSe (Ref. 40); that for ZnS should be very similar. For the purpose of discussion the band structure has been shown divided into three energy regions, labelled 1–3; also shown is an energy threshold $\xi_a \sim 2.2$ eV for impact excitation of the Mn^{2+} activator. Electrons with energy above this threshold can be inelastically scattered to a lower energy region of the band by collision with Mn^{2+}. The significance of these energy regions is as follows:

(i) Region 1 is essentially the central valley with the CB mimimum at Γ.

(ii) Region 2 includes the low energy satellite valleys at the L and X points. In this region the bands are comparatively flat over a large part of the Brillouin zone; electrons in this energy region have low velocity and can be represented by

4566   J. Appl. Phys., Vol. 54. No. 8, August 1983

57

Robbins et al   4566

localized wave functions constructed from the Bloch waves. Electrons with energy greater than the threshold ξ_a in region 2 therefore have a relatively high probability of impact interaction with the localized Mn^{2+} centers.

(iii) Region 3 includes all the higher bands.

Consider electrons entering a high field region, with energy initially near the Γ minimum in region 1. The electrons are accelerated in the central valley, the dominant energy loss process being inelastic polar optical scattering.[41] The efficiency of the latter process decreases as soon as the electron energy exceeds a few times the optical phonon energy, so that the electrons are quite rapidly heated in the central valley and the "temperature" of the distribution, T_e, increases sharply with field strength. At an energy corresponding to region 2, intervalley scattering becomes dominant and transfer to the satellite valleys occurs. The strong intervalley scattering ensures that the electrons in region 2 have a Maxwellian distribution with T_e close to the lattice temperature, although the overall electron distribution is very non-Maxwellian. The average energy of electrons in the satellite valleys does not increase very rapidly with field strength.[42] Simulations of transfer to the satellite valleys in GaAs devices[42,43] suggest that for fields approximately $10\times$ the threshold field the majority of electrons will have transferred. Applying the same criterion to these II-VI compounds suggests that for fields $\gtrsim 0.4$ MV cm^{-1} in ZnSe the majority of electrons will transfer to region 2; ZnS should behave similarly. At a field of 0.4 MV cm^{-1} the transition to region 2 could occur in a transit length of ~ 400 Å. The slow rate of heating with increasing field strength in region 2 makes it difficult for electrons to escape to higher energy. However, when the field is sufficiently high that some electrons do reach region 3, the low density-of-states at energies just above region 2 allows these electrons to be rapidly accelerated to energies exceeding the threshold for lattice ionization. For this reason the onset of avalanching may be a very sharp function of increasing field, as proposed by Howard *et al.*[44] in their discussion of the "memory" effect in ac thin film electroluminescent devices.

The most likely explanation of the constant power efficiency represented by Eq. (44) is, therefore, that in the field range $0.6 < \bar{E}_z < 1.2$ MV cm^{-1} the majority of electrons in the active part of the device occupy region 2 in the electron energy space. Assuming that electrons lose energy only by creation of phonons or by impact excitation of the activator, the power efficiency can be expressed as an energy loss ratio:

Power efficiency

$$= \frac{\text{rate of energy loss to activator}}{\text{total rate of energy loss}}. \quad (45a)$$

Neglecting contributions from regions 1 and 3, this relationship can be expressed crudely as

$$\frac{B}{I\bar{E}_z} \propto \frac{\bar{\phi}_a \bar{\xi}_a F_a}{[\bar{\phi}_p \bar{\xi}_p + \bar{\phi}_a \bar{\xi}_a F_a]}, \quad (45b)$$

where $\bar{\phi}_p$ and $\bar{\phi}_a$ are average rates per unit length for phonon creation and activator excitation appropriate to region 2, and $\bar{\xi}_p$ and $\bar{\xi}_a$ the corresponding average energies. F_a is the fraction of electrons in region 2 with an energy above threshold for impact excitation of the activator. It was pointed out above that the strong intervalley scattering in region 2 means that the average energy of electrons in this region does not increase very fast with increasing field. Hence there should exist a field range, just above the field required for efficient population of the L and X satellite valleys, in which the fraction F_a changes only slowly with increasing field, i.e., where the electron distribution is Maxwellian with a temperature close to the lattice temperature. In that field range the quantity on the right-hand side of expression (45b) would be roughly constant, and in accordance with the experimental observations expressed in Eq. (44).

At sufficiently high fields some electrons from the upper part of region 2 will be accelerated into region 3, where they may begin to lose energy by lattice ionization. Using an average rate and energy for ionization, $\bar{\phi}_i$ and $\bar{\xi}_i$, expression (45b) in this high field range can be simply modified to give

$$\frac{B}{I\bar{E}_z} \propto \frac{\bar{\phi}_a \bar{\xi}_a F_a'}{[\bar{\phi}_p \bar{\xi}_p + \bar{\phi}_a \bar{\xi}_a F_a' + \bar{\phi}_i \bar{\xi}_i F_i']}, \quad (46)$$

where F_a' and F_i' represent the fractions of the total electron population in regions 2 and 3 with energies above threshold for activator excitation and for lattice ionization, respectively. The fraction F_i' will increase much more rapidly with increasing field above threshold for ionization than will F_a'. Therefore, the additional loss term appearing in the denominator of Eq. (46), compared with Eq. (45), may cause a decrease in QI $(= B/I)$ as soon as the threshold field for lattice ionization is exceeded.

The valence band holes generated by lattice ionization in the high field region can contribute to the positive space charge accumulating in the ZnS:Mn layer. However, this space charge itself tends to enhance the field in the ZnS:Mn layer nearer the cathode, and to reduce it in the layer nearer the anode (gate). This correlation between high electric field and additional space-charge generation should produce large field distortion in the ZnS:Mn layer. When ionization becomes significant in forward bias there will be a tendency for an increasingly high field region to be constrained to a steadily decreasing length of the ZnS:Mn layer near the SiO$_2$ interface; at the same time the field in the remainder of the ZnS:Mn layer will tend to fall. Initially, when the high field region extends through an appreciable fraction of the ZnS:Mn layer thickness, those high energy electrons which lose energy by lattice ionization may not contribute to the Mn^{2+} luminescence simply because there is insufficient transit length remaining in the layer to reheat them to the activator impact threshold. This will reduce the quantum efficiency of the luminescence. However, as the accumulating space charge restricts the increasingly high field region to shorter and shorter distances in the ZnS:Mn layer, even those electrons which cause ionization and emerge from the high field region with energies low in region 1 of the electron energy space may have a high probability of being reheated and causing impact excitation of the activator in the remaining length of the layer, providing the field in the bulk is maintained above threshold for this process. As a result the quan-

4567   J. Appl. Phys., Vol. 54, No. 8, August 1983

58

Robbins *et al.*   4567

tum efficiency will tend to rise again towards a value characteristic of the medium field regime, i.e., that pertaining before lattice ionization becomes significant. These effects, deriving from the field distortion in the ZnS:Mn layer, may be the explanation of the dip in the QI values evident in Fig. 12. If so, this suggests that ionization of the ZnS:Mn layer becomes important when the average field in the oxide layer $\bar{E}_0 \sim 3.4$ MV cm^{-1}. Since electron trapping is only just becoming significant at this field, the field $E(l_z^-)$ in the oxide at the SiO$_2$/ZnS:Mn interface will be close to this value. Allowing for the difference in dielectric constants, and assuming there is initially a negligible concentration of positive space charge right at the interface, suggests an onset of ionization at a field ~ 1.7 MV cm^{-1} in the ZnS:Mn layer. This same value for the ionization threshold has been proposed in order to explain the hysteretic "memory" phenomenon in ACTFEL devices.[44]

E. Summary and conclusions

The explanation of the experimental results presented in Figs. 12 and 13 can be summarized in the following way. When the average field in the ZnS:Mn layer is in the range $0.6 < \bar{E}_z < 1.2$ MV cm^{-1}, and when the field distortion arising from space-charge accumulation is small, the luminescence intensity per unit current is proportional to the average field, i.e., the light output is simply proportional to the average power dissipated in the active layer. In fields of this magnitude electrons are quite rapidly heated to reach the satellite L and X CB minima, region 2 of Fig. 15, where they become "trapped" by the strong intervalley scattering which slows the rate of further heating. The hot electrons accumulate in this energy region. The threshold for impact excitation of Mn^{2+} occurs at an energy within region 2, so that some fraction of the hot electron distribution (which changes only slowly with field) can excite the activator. In fields $\bar{E}_z > 0.6$MV cm^{-1} the electron transit length to reach the impact threshold is quite short (< 400 Å); a quasi-steady state is rapidly established as the hot electron distribution drifts down the remaining length of the active layer. Efficient heating to region 2 in ZnS would limit the advantage of "hot" electron injection over the potential step at the SiO$_2$/ZnS interface to very thin active layers, or to layers with low average fields, and in practice such "hot" injection is difficult to engineer because of the intrinsically poor quality of the first-to-grow ZnS on SiO$_2$.

When the local field exceeds ~ 1.7 MV cm^{-1}, some electrons are excited into region 3 and are rapidly heated to the ionization threshold. When the high field region extends over an appreciable length of the active layer, the loss of electron energy due to ionization causes a fall in quantum efficiency for the Mn^{2+} luminescence. However, accumulation of positive space charge near the SiO$_2$ interface gradually reduces the penetration of the high field into the active layer, so that the medium field regime (< 1.7 MV cm^{-1}) favoring impact excitation is reestablished over most of the length of the ZnS:Mn layer. Ionization is then restricted to a thin layer near the SiO$_2$ interface. The steady-state accumulation of electrons in the satellite valleys, and subsequent

drift, can occur again over most of the active layer length, producing a rise in quantum efficiency to its previous value. Eventually the accumulation of positive space charge reduces the average field $\bar{E}_z$ below threshold for maintaining the electron population in the satellite valleys, and the Mn^{2+} luminescence efficiency falls. However, the data in Fig. 12 suggest that the decrease in the steady-state hot electron population of energy region 2 does not become important until the average field $\bar{E}_z$ falls below ~ 0.4 MV cm^{-1}.

The peak external power efficiency of the devices studied here is $\sim 8 \times 10^{-4}$ W/W; this value is reduced from the excitation efficiency by a combination of light trapping, absorption in the Si substrate and in the ITO gate electrode, and by the radiative efficiency η_r of the Mn^{2+} activator if this is less than unity. A reasonable estimate of the excitation efficiency is probably $\sim 10^{-2}$W/W in our devices. The right-hand side of expression (45) represents a crude upper limit to the power efficiency, and since this is < 1 we have

$$\eta_{power}^{max} \sim \frac{\bar{\phi}_a \bar{\xi}_a F_a}{\bar{\phi}_p \bar{\xi}_p} \sim 10^{-2}. \tag{47}$$

For ZnS:Mn, $\bar{\xi}_a \sim 2.2$ eV, and assuming energy loss to optical phonons is dominant, $\bar{\xi}_p \sim 0.04$ eV. The impact excitation rate per unit length is given by

$$\bar{\phi}_a \sim \sigma N_a. \tag{48}$$

The cross section has been recently estimated[45] as $\sim 2 \times 10^{-16}$ cm^{-2}, which is a geometric cross section and likely to be an upper limit for a spin-forbidden excitation of a neutral impurity[33]; the optimum EL efficiency typically occurs for a Mn^{2+} concentration $N_a \sim 1.5 \times 10^{20}$ cm^{-3}. These values give $\bar{\phi}_a \lesssim 3 \times 10^4$ cm^{-1}. An earlier analysis[38] of ionization in ZnSe according to Baraff's theory suggests a phonon scattering length in the range 25–52 Å; using this as an estimate gives $\bar{\phi}_p \sim 3 \times 10^6$ cm^{-1}. Hence, substitution in Eq. (47) gives a lower limit for the fraction F_a:

$$F_a \gtrsim 0.02. \tag{49}$$

The fraction is smaller than that derived by Tornquist and Tuomi,[30] although the uncertainties in all the estimated quantities are too large to draw firm conclusions. However, the general approach appears reasonable, and efforts to improve the quantitative estimates of cross sections and scattering rates should result in better understanding of the factors limiting EL efficiency.

It is well known that Mn^{2+} is the most efficient activator in this kind of high-field electroluminescent device. Its advantages include good solubility in the host matrix, and high intrinsic radiative and luminous efficiencies; in addition the luminescent divalent state with the half-filled shell configuration is very stable against change in oxidation state. However, the most important factor favoring Mn^{2+} may be the relatively low impact threshold energy, which allows excitation by a significant fraction of the hot electron distribution suggested to accumulate in region 2 of Fig. 15. The other activator widely studied in ZnS, Tb^{3+}, has an efficiency significantly lower than that of Mn^{2+}. The lowest energy $^7F_6 \rightarrow ^5D_4$ excitation occurs at an energy ~ 2.6 eV, i.e., ~ 4

eV higher than the $^6A_1 \rightarrow {^4}T_1$ excitation of Mn^{2+}. The lower efficiency of ZnS:Tb may, therefore, result largely from the fact that a smaller fraction of the steady-state hot electron distribution in region 2 occurs at an energy above threshold for the Tb^{3+} activator.

Finally, we note that measurements of the kind reported here are clearly valuable in the study of space-charge accumulation and annihilation in the ZnS:Mn layer. Such measurements might be relevant to the study of "memory" effects in ACTFEL layers of higher Mn concentration, which are believed to be determined by space-charge effects.[14]

ACKNOWLEDGMENTS

[1]D. J. Robbins, C. Falcony, D. J. DiMaria, D. W. Dong, J. F. DeGelormo, I. F. Chang, and D. B. Dove, IEEE Electron Dev. Lett. EDL-3, 148 (1982).

[2]D. J. DiMaria, in *The Physics of MOS Insulators*, edited by G. Lucovsky, S. T. Pantelides and F. L. Galeener (Pergamon, New York 1980), p. 1.

[3]D. Langer and S. Ibuki, Phys. Rev. 138, A809 (1965).

[4]T. Inoguchi, M. Takeda, Y. Kakihara, Y. Nakata, and M. Yoshida, SID Dig. 74, 84 (1974).

[5]J. Kirton, in *Handbook on Semiconductors*, series edited by T. S. Moss, Vol. 4 edited by C. Hilsum (North-Holland, Amsterdam, 1981).

[6]J. W. Allen, J. Lumin. 7, 228 (1973).

[7]M. I. Abdalla, J. Thomas, A. Brenac, and J-P. Noblanc, IEEE Trans.

tron Devices ED-28, 694 (1981).

[8]H. Ohnishi and Y. Hamakawa, Jpn. J. Appl. Phys. 17, 1225 (1978).

[9]T. Mishima, W. Quan-kun, and K. Takahashi, J. Appl. Phys. 52, 5797 (1981).

[10]J. J. Hanak, Jpn. J. Appl. Phys. Suppl. 2, 809 (1974).

[11]C. J. Alder, A. F. Fray, C. Hilsum, and P. Lloyd, Displays 1, 191 (1980).

[12]D. J. DiMaria, R. Ghez, and D. W. Dong, J. Appl. Phys. 51, 4830 (1980).

[13]J. A. Cape, R. D. Ketchpel, and L. G. Hale, IEEE Trans. Electron Devices ED-25, 1352 (1978).

[14]W. E. Howard, J.Lumin. 23, 155 (1981); D. H. Smith, *ibid.* 23, 209 (1981).

[15]D. C. Morton and F. E. Williams, SID Dig. 81, 30 (1981).

[16]D. Dong, E. A. Irene, and D. R. Young, J. Electrochem. Soc. 128, 819 (1978).

[17]C. A. Pan and T. P. Ma, Appl. Phys. Lett. 37, 714 (1980).

[18]M. Lenzlinger and E. H. Snow, J. Appl. Phys. 40, 278 (1969).

[19]D. J. DiMaria and D. W. Dong, J. Appl. Phys. 51, 2722 (1980).

[20]S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).

[21]D. J. Robbins, C. Falcony, and D. J. DiMaria (unpublished work).

[22]W. Reuter (private communication).

[23]V. Marrello, L. Samuelson, A. Onton, and W. Reuter, J. Appl. Phys. 52, 3590 (1981).

[24]Z. A. Weinberg, W. C. Johnson, and M. A. Lampert, Appl. Phys. Lett. 25, 42 (1974).

[25]R. Chin, N. Holonyak, Jr., G. E. Stillman, J. Y. Tang, and K. Hess, Electron. Lett. 16, 467 (1980).

[26]H. Venghaus, D. Theis, H. Oppolzer, and S. Schild, J. Appl. Phys. 53, 4146 (1982).

[27]A. F. Cattell and A. G. Cullis, Thin Solid Films 92, 211 (1982).

[28]H. Ohnishi and Y. Hamakawa, Jpn. J. Appl. Phys. 16, 981 (1977).

[29]H. Matsumoto, S. Tanaka, and T. Yabumoto, Jpn. J. Appl. Phys. 17, 1543 (1978).

[30]R. O. Tornquist and T. O. Tuomi, J. Lumin 24/25, 901 (1981).

[31]N. T. Gordon, M. D. Ryall, and J. W. Allen, Appl. Phys. Lett. 35, 691 (1979).

[32]N. T. Gordon, IEEE Trans. Electron Devices ED-28, 434 (1981).

[33]J. W. Allen, J. Lumin. 7, 228 (1973).

[34]J. W. Allen, J. Lumin. 23, 127 (1981) (and other papers in this volume).

[35]D. C. Krupka, J. Appl. Phys. 43, 476 (1972).

[36]H. Kobayashi, S. Tanaka, and H. Sasakura, Jpn. J. Appl. Phys. 12, 1637 (1973).

[37]G. W. Ludwig and M. Aven, J. Appl. Phys. 38, 5326 (1967).

[38]A. W. Livingstone and J. W. Allen, J. Phys. C 3, 2468 (1970).

[39]N. T. Gordon and J. W. Allen, Solid State Commun. 37, 441 (1981).

[40]J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B 14, 556 (1976).

[41]E. M. Conwell, *High Field Transport in Semiconductors*, Solid State Physics, Suppl. 9 (Academic, New York, 1967).

[42]W. Fawcett, A. D. Boardman, and S. Swain, J. Phys. Chem. Solids 31, 1963 (1970).

[43]D. Jones and H. D. Rees, J. Phys. C 6, 1781 (1973).

[44]W. E. Howard, O. Sahni, and P. M. Alt, J. Appl. Phys. 53, 639 (1982).

[45]J. E. Bernard, M. F. Martens, and F. Williams, J. Lumin. 24/25, 893 (1981).

Characterization of Plasma-Enhanced Chemically-Vapor-Deposited

Silicon-Rich Silicon Dioxide/Thermal Silicon Dioxide

Dual Dielectric Systems

S. Yokoyama[a], D. W. Dong, D. J. DiMaria, and S. K. Lai[b]

I.B.M. Thomas J. Watson Research Center

Yorktown Heights, New York 10598

ABSTRACT

Plasma enhanced chemically-vapor-deposited silicon-rich oxides (200 Å and 500 Å in thickness) of various excess silicon content were deposited onto thermal silicon dioxide (SiO_2) layers (103, 207 and 530 Å in thickness) grown on a p-type silicon (Si) substrate. The dielectric constant, electron injection efficiency, current-voltage (I-V) reproducibility and breakdown property of these composite structures were examined. The dielectric constants of Si-rich oxide were observed to increase with Si content from 3.8 for films deposited at a gas phase ratio (R_0) of the concentration of nitrous oxide (N_2O) to silane (SiH_4) of 150 to ~10 for films deposited with $R_0=0$. The Si-rich oxides with $R_0 \leq 5$ were found to work as electron injectors. The average oxide field needed to induce a current of 4.8×10^{-7} A/cm^2 through the SiO_2 (530 Å in thickness) decreased about 40% in magnitude by adding a Si-rich oxide layer with the optimized R_0 ($=1$) compared to that of a control sample which had no Si-rich oxide layer. For thin SiO_2 (103 Å and 207 Å in thickness) samples, the decrease of the average field was only 2% and 10% in magnitude with the optimized R_0 ($=2$) layer, respectively, due to the relatively large voltage drop (≈ -1 V) across the Si-rich

oxide compared to that across the thermal oxide layer. The voltage drop across the oxide is discussed in terms of a dual dielectric model. The yield, which was defined as the percentage of capacitors that required a field larger than 2 MV/cm to obtain a current of 9.6×10^{-4} A/cm^2, on as-fabricated samples was larger than 90% for all samples with Si-rich oxide. The samples were not destroyed by the passage of a relatively high current density (1.21×10^{-2} A/cm^2) through the oxide and subsequent measurements resulted in approximately the same field to produce the specified current as for the first measurement. The yield was found to have a maximum at $R_0 = 1\text{-}10$ depending on the thicknesses of Si-rich oxides and SiO$_2$. Current-voltage reproducibility was also found to be improved by the deposition of Si-rich oxide.

I. Introduction

Recently, Si-rich oxides with an excess Si content of about 13% made by Atmospheric-Pressure Chemical-Vapor-Deposition (APCVD) have been used as electron injectors for Electrically-Alterable Read-Only-Memories (EAROMs) operated at low voltages (1-3) and for yield improvement on Metal-Oxide-Semiconductor (MOS) structures with very thin oxides (4). Slightly off-stoichiometric oxides (1% to 6% excess silicon) are also used for intervening layers in more recent EAROMs with extended cyclability (3). The physical mechanism of the current conduction through Si-rich oxide films has also recently been studied (5).

Various characteristics of Si-rich oxide (5-7) such as enhanced electron injection, high capacitance, yield improvement, and low permanent trapping are due to the special microstructure of the Si-rich oxide which has at least two phases (Si and SiO_2) (8-11). The shape irregularity of the Si islands in the SiO_2 matrix causes the field enhancement at the Si-rich oxide/SiO_2 interface. One of the great advantages of the Si-rich oxide is that both (top and bottom) surfaces of the film can act as enhanced electron injection interfaces with SiO_2. This is different from polycrystalline silicon (poly-Si) films in which only the top surface gives enhanced electron injection (12-14). This special feature of the Si-rich oxide is necessary for making Dual-Electron-Injector-Structures (DEISs) which are used in EAROMs (2).

Insulators like silicon nitride (15) and tantalum pentoxide (16) have been proposed and used to give higher dielectric constants compared to that of thermal oxide. The use of such insulators (in particular, tantalum pentoxide) is still experimental and their compatibility with standard poly-Si gate processing and long term

reliability are questionable. However, the use of Si-rich oxide is only a minor pertur-bation in normal poly-Si gate processing. Another advantage of Si-rich oxide is the very low permanent charge trapping in the film. These characteristics of the Si-rich oxide in a composite stack with SiO_2 can be used in high-capacitance memory struc-tures where high dielectric constants and little permanent charge trapping are desira-ble (4).

Previous work has utilized APCVD Si-rich oxide films deposited at ~700 °C with excess Si content up to $\approx 13\%$ (1-7,9-11,17). However, the detailed relation between the excess Si content and the characteristics of the Si-rich oxide films used in composite structures as previously mentioned has not been extensively studied. It is expected that there is a optimum excess Si content. A film with 100% atomic Si content should be similar to poly-Si, while a film with 0% excess atomic Si content should be SiO_2. It is important to determine the optimum excess Si content for application of the Si-rich oxide films in actual devices. Furthermore, from the correlation between electronic and structural properties of the Si-rich oxide films for various excess Si contents, the physical mechanisms operative in these films can be more clearly understood.

In this paper the detailed relation between the gas phase ratio R_0 defined as the ratio of the concentration of N_2O to SiH_4, which affects the excess Si content in the Si-rich oxide, and the electronic properties such as electron injection efficiencies, current-voltage (I-V) reproducibility, and breakdown yield will be investigated.

II. Sample Preparation

Plasma-Enhanced Chemical-Vapor-Deposition (PECVD) was used for the deposition of the Si-rich oxide. This method has the advantage of low temperature deposition and good uniformity of the film thickness. Furthermore, it is easier with PECVD to change the excess Si content in the films from poly-Si to stoichiometric SiO_2 as compared to thermal APCVD (17).

Si-rich oxide was deposited by using a capacitively coupled rf discharge system (Plasma-therm Model PK 1250). SiH_4 (2%, diluted with Ar) and N_2O (100%) gases were utilized. The gas phase ratio, $R_0=[N_2O]/[SiH_4]$, was varied from 0 to 150 to change the excess Si content in the Si-rich oxide. For $R_0 \leq 45$, the SiH_4 flow rate was kept constant (500 standard-cubic-centimeter-per-minute [sccm]) and the N_2O flow rate was changed. At $R_0=150$, the SiH_4 flow rate was decreased to 200 sccm due to the limit of the pumping speed of the vacuum system. Substrate temperature, total pressure, rf (13.56 MHz) power, and the electrode distance were, respectively, kept at 350°C, 80 Pascal (Pa), 0.25 W/cm^2, and 5 cm. High temperature annealing was done for some samples after deposition of Si-rich oxide in an N_2 ambient at 1000°C for 30 minutes. Post metallization annealing was done in forming gas (90% N_2 and 10% H_2) at 400°C for 20 min. After gate metallization of aluminum, Si-rich oxide layers with $R_0 \leq 3$ were etched off by Reactive-Ion-Etching (RIE) using the aluminum gate as an etching mask. The gate areas investigated were 9.23×10^{-3} cm^2, 5.19×10^{-3} cm^2, 2.07×10^{-3} cm^2 and 4.13×10^{-4} cm^2.

The sample configuration used here which has various thicknesses of Si-rich oxide and thermal oxide is shown in Figure 1. This structure is called a Single-

65

Electron-Injector-Structure (SEIS), as compared to the DEIS previously discussed. For the thin (103 and 207 Å in thickness) thermal oxide samples, a thick field oxide (3500 Å) was applied outside the gate area to protect the edge region. The surface morphology of the structures was studied using Scanning-Electron-Microscopy (SEM), and the results will be discussed briefly.

III. Measurements

Dark current-voltage measurements were done either at a constant gate voltage ramp rate using a Keithley #26000 logarithmic picoammeter and a voltage source with adjustable ramp rates, or point-by-point using a voltage stepping unit. The small-signal (15 mV) ac-capacitance as a function of gate voltage (C-V) was measured at a frequency of 1 MHz using a Boonton capacitance meter (model 72BD). All measurements were performed at room temperature.

IV. Results and Discussion

The samples discussed in this section are SEISs discussed in Section II and ordinary MOS structures for control samples.

A. Current conduction and dielectric constant of "bulk" Si-rich SiO_2

To facilitate in understanding the properties of the Si-rich SiO_2 materials, a series of samples were made which use only this material as the insulator without the intervening SiO_2 layer. Figure 2 shows typical I-V characteristics of Al/Si-rich oxide (425 to 600 Å in thickness)/p-Si structures with different amounts of excess Si. These samples are all annealed at high temperature ($1000°C-N_2-30$ min). With

II. Sample Preparation

Plasma-Enhanced Chemical-Vapor-Deposition (PECVD) was used for the deposition of the Si-rich oxide. This method has the advantage of low temperature deposition and good uniformity of the film thickness. Furthermore, it is easier with PECVD to change the excess Si content in the films from poly-Si to stoichiometric SiO_2 as compared to thermal APCVD (17).

Si-rich oxide was deposited by using a capacitively coupled rf discharge system (Plasma-therm Model PK 1250). SiH_4 (2%, diluted with Ar) and N_2O (100%) gases were utilized. The gas phase ratio, $R_0=[N_2O]/[SiH_4]$, was varied from 0 to 150 to change the excess Si content in the Si-rich oxide. For $R_0 \leq 45$, the SiH_4 flow rate was kept constant (500 standard-cubic-centimeter-per-minute [sccm]) and the N_2O flow rate was changed. At $R_0=150$, the SiH_4 flow rate was decreased to 200 sccm due to the limit of the pumping speed of the vacuum system. Substrate temperature, total pressure, rf (13.56 MHz) power, and the electrode distance were, respectively, kept at 350°C, 80 Pascal (Pa), 0.25 W/cm^2, and 5 cm. High temperature annealing was done for some samples after deposition of Si-rich oxide in an N_2 ambient at 1000°C for 30 minutes. Post metallization annealing was done in forming gas (90% N_2 and 10% H_2) at 400°C for 20 min. After gate metallization of aluminum, Si-rich oxide layers with $R_0 \leq 3$ were etched off by Reactive-Ion-Etching (RIE) using the aluminum gate as an etching mask. The gate areas investigated were 9.23×10^{-3} cm^2, 5.19×10^{-3} cm^2, 2.07×10^{-3} cm^2 and 4.13×10^{-4} cm^2.

The sample configuration used here which has various thicknesses of Si-rich oxide and thermal oxide is shown in Figure 1. This structure is called a Single-

Electron-Injector-Structure (SEIS), as compared to the DEIS previously discussed. For the thin (103 and 207 Å in thickness) thermal oxide samples, a thick field oxide (3500 Å) was applied outside the gate area to protect the edge region. The surface morphology of the structures was studied using Scanning-Electron-Microscopy (SEM), and the results will be discussed briefly.

III. Measurements

Dark current-voltage measurements were done either at a constant gate voltage ramp rate using a Keithley #26000 logarithmic picoammeter and a voltage source with adjustable ramp rates, or point-by-point using a voltage stepping unit. The small-signal (15 mV) ac-capacitance as a function of gate voltage (C-V) was measured at a frequency of 1 MHz using a Boonton capacitance meter (model 72BD). All measurements were performed at room temperature.

IV. Results and Discussion

The samples discussed in this section are SEISs discussed in Section II and ordinary MOS structures for control samples.

A. Current conduction and dielectric constant of "bulk" Si-rich SiO_2

To facilitate in understanding the properties of the Si-rich SiO_2 materials, a series of samples were made which use only this material as the insulator without the intervening SiO_2 layer. Figure 2 shows typical I-V characteristics of Al/Si-rich oxide (425 to 600 Å in thickness)/p-Si structures with different amounts of excess Si. These samples are all annealed at high temperature ($1000°C-N_2-30$ min). With

increasing R_0, the resistivity is increased. The hysterisis of the I-V curve is drastically increased at $R_0 = 10$. The abnormal behavior of the resistivity in thermal APCVD films previously reported (17), with the resistivity of the film with $R_0 < 3$ larger than that of the film with $R_0 = 3$, was not observed. In Fig. 3, the I-V curves were fitted to the expression for planar Fowler-Nordheim tunneling where the pre-exponential field dependence has been neglected (18-20). The effective barrier heights are shown in Fig. 4 as a function of R_0. With increasing excess Si content in the films (decreasing R_0), the effective barrier heights decreased due to a decrease in the tunneling distance between Si islands (5).

Figure 5 shows the C-V curves at a frequency of 1 MHz for Al/Si-rich oxide (200 Å in thickness)/thermal SiO_2 (530 Å in thickness)/p-Si structures with different values of R_0 for the Si-rich oxide. The accumulation capacitance is decreasing with increasing R_0. Figure 6 shows the effective dielectric constant of the Si-rich oxide as a function of R_0 derived from the accumulation capacitance of the C-V curves shown in Fig. 5. The scattering of the data is due to the non-uniformity of the Si-rich oxide thickness, deviation in the gate areas, and the fluctuation of the film conductance (effective dielectric constant was determined by assuming that the Si-rich oxide is a perfect insulator). The effective dielectric constant increases as R_0 is decreased. The contribution of the non-uniformity of the Si-rich oxide thickness is consistent with the observation that the scatter in the data for composites using thick (500 Å) Si-rich oxide layers is smaller than that for structures using thin (200 Å) Si-rich oxide layers. Effective dielectric constants larger than 6 are obtained for $R_0 \leq 3$.

B. Electron injection efficiency and I-V reproducibility

Figure 7 shows typical I-V curves for the Al/Si-rich oxide (200 Å in thickness)/SiO$_2$ (530 Å in thickness)/p-Si structures with different R$_0$ where the applied voltage was ramped at -0.5 V/sec from 0 V. The return traces were omitted for clarity. The Si-rich oxides with R$_0 \leq 5$ were found to work as enhanced electron injectors. The I-V curves for R$_0 \leq 5$ are all located at lower negative gate voltages than that of the control sample which has no Si-rich oxide layer. For R$_0 \geq 10$, the films have a high resistivity, consistent with Figs. 2 and 4. For the control and high resistivity samples with R$_0 \geq 10$, destructive voltage breakdown occured before the current exceeded 5 μA (9.6×10^{-4} A/cm^2). However, the samples with R$_0 \leq 5$ Si-rich oxide layers were not broken-down destructively, even for currents exceeding 5 μA. For the second and subsequent measurements, the I-V curves were shifted to higher negative voltages by the permanent charge trapping in SiO$_2$ layer. The Si-rich oxide films with R$_0 = 0$ should behave like a poly-Si gate (21). For these samples, the Si-rich oxide layer still acts as an enhanced electron injector. This could be attributed either to the rougher interface between these films deposited by PECVD and thermal oxide or to the possible porous nature of the thin Si-rich oxide films where the Al metal counter electrode could penetrate in fingerlike extensions to the SiO$_2$ interface. SEM photographs of these films did not show any porous structure within a resolution of ~200 Å.

Figure 8 shows the gate voltage at 1 nA (1.92×10^{-7} A/cm^2) for Al/Si-rich oxide (200 Å in thickness)/SiO$_2$ (530 Å in thickness)/p-Si structures. Similar results were obtained for 500 Å Si-rich oxide composite structures. The gate voltages for samples annealed at high temperature (1000°C-N$_2$-30 min) are smaller than for

samples that did not have high temperature annealing. This might imply that Si grains in the Si-rich oxide were enlarged or made rougher during the high temperature annealing which resulted in a higher field enhancement. Another possible explanation for this could be that the Si-rich oxide films were made more conductive with annealing and had a smaller amount of the total voltage dropped across their thickness. This will be discussed later. For APCVD Si-rich oxides with $R_0=3$, high temperature annealing does not affect the injection efficiency very much (6), consistent with the observations near $R_0=3$ for the PECVD films of Fig. 8.

Figure 9 shows the I-V reproducibility histograms for 5 μA (9.6×10^{-4} A/cm^2) currents on the same series of samples as shown in Fig. 7. For the control sample, this histogram shows mostly the destructive breakdown field, not I-V reproducibility. For the samples with $R_0 \leq 5$ Si-rich oxide layers, almost the same histogram was obtained with a second measurement. Permanent charge trapping in the SiO_2 layer during the first ramp will move the histograms observed during the second ramp to higher electric fields. The electric field at the peak of the histogram was decreased by adding the Si-rich injectors as compared to the control sample. The average electric field used in Fig. 9 and subsequent figures is calculated by dividing the magnitude of the gate voltage by the SiO_2 layer thickness after corrections of at most ≈ 1 V in magnitude for the gate metal-Si substrate work function difference and the silicon surface potential. No correction for the finite voltage dropped across the Si-rich oxide layer has been made, and this will be discussed in detail later. I-V reproducibility is also improved by the Si-rich oxide layer, and it was found to be best at $R_0=3$. Figure 10 shows the gate voltage at the peak of each histogram as a function of R_0. For the control sample, this plot shows the destructive breakdown

voltage. The error bar indicates the half width of the distribution for the histogram. Even taking into account the error bar, the sample with $R_0=1$ still has a minimum gate voltage, implying highest electron injection efficiency. Similar results were obtained for the Al/Si-rich oxide (500 Å in thickness)/SiO$_2$ (530 Å in thickness)/p-Si structures.

Samples with thinner SiO$_2$ layers were prepared to examine the possible use of PECVD Si-rich oxides in very large scale integration (VLSI). These samples have a thick field oxide (3500 Å in thickness) surrounding the gate area to protect the edge region of thin SiO$_2$ films, and have PECVD Si-rich oxide layers (200 Å in thickness), with various R_0, deposited on the thin SiO$_2$ layer. Figure 11 shows the average electric field required to produce a current of 1 nA (4.83×10^{-7} A/cm^2) for the structures with 103, 207, and 530 Å thick SiO$_2$ layers. Although for the 530 Å sample the average field is decreased about 40% in magnitude by adding a Si-rich oxide injector compared to that of the control sample, for thinner oxide samples (103 and 207 Å in thicknesses) the decrease of the average field from that of control samples are only 2% and 10% for 103 and 207 Å SiO$_2$ samples, respectively.

These differences between the structures are believed to be due to the voltage drop across the Si-rich oxide injector which has been ignored in previous discussions. The voltage drop across the Si-rich oxide layer, V_n, can be shown from electrostatics to be (7),

$$V_n = [(V_g - \Phi_{ms} - \Psi_s) + \frac{qN_0(l_n + l_0 - \bar{x}_0)}{\epsilon_0} + \frac{qN_n\bar{x}_nl_0}{\epsilon_0l_n}]x[1 + \frac{\epsilon_nl_0}{\epsilon_0l_n}]^{-1}$$ (1)

for a dual dielectric system. Here, q is the magnitude of the charge on an electron (1.6×10^{-19} coul), V_g is the gate voltage, l_0 (l_n) is the thickness of the SiO_2 layer (Si-rich oxide layer), Φ_{ms} is the work function difference between the metal electrode and semiconductor substrate (by convention, expressed in volts), Ψ_s is the silicon surface potential, $\bar{x}_0$ ($\bar{x}_n$) is the centroid of the trapped charge distribution in the SiO_2 layer (Si-rich oxide layer) as measured from the Al interface with the Si-rich oxide, N_0 (N_n) is the concentration of the "bulk" trapped electrons per unit area of this distribution in the oxide layer (Si-rich oxide layer), and ϵ_0 (ϵ_n) is the low frequency permitivity of SiO_2 (Si-rich oxide). N_0 can be assumed to be zero since at low fields (≤ 4 MV/cm) the current injection into the SiO_2 is low enough so that no significant trapped SiO_2 charge will build up (7), and $\bar{x}_n$ is assumed to be approximately l_n. Then equation (1) reduces to

$$V_n = [(V_g - \Phi_{ms} - \Psi_s) + \frac{qN_n l_0}{\epsilon_0}]x[1 + \frac{\epsilon_n l_0}{\epsilon_0 l_n}]^{-1}. \tag{2}$$

Determination of N_n is difficult since a reversible space charge builds up in the Si-rich oxide layer. The build-up of this charge is demonstrated in Fig. 12 where a point-by-point I-V measurement has been used to minimize capacitive effects and allow studies of low currents ($\leq 10^{-10}$ A). The apparent current ledge observed during the "up" trace from 0 V to -30 V at low voltages in this figure is due to the space charge build up in the Si-rich oxide injector layer (6). This tends to influence and hold-off current injection from the contact until a voltage is reached (≈ -23 V for the structure and conditions of Fig. 12) where particle current injection into the SiO_2 layer controls the measured current in the external circuit. The current reversal in this figure during the "down" trace at low voltages is due to this reversible space

73

charge in the Si-rich oxide layer flowing back into the gate contact due to its own internal field (6). For thinner Si-rich SiO_2 layers or layers with a higher conductivity (more excess Si), this space charge would build up much more rapidly and push the apparent ledge to even lower current levels.

Reasonable estimates of N_n would be in the range from 10^{12} to 10^{13} electrons/cm^2. Substituting these values and $V_g = -8.7$ V at 1 nA (4.83×10^{-7} A/cm^2), $\Phi_{ms} = -0.91$ eV, $\Psi_s = -0.16$ eV, $\epsilon_n = 9.3 \times (8.86 \times 10^{-14}$ F/cm), $\epsilon_0 = 3.9 \times (8.86 \times 10^{-14}$ F/cm), and the appropriate insulator thickness for the Al/Si-rich SiO_2 (200 Å in thickness, $R_0 = 1$)/SiO_2 (103 Å in thickness)/p-Si structure in equation (2) yields values of $V_n = -3.2$ V and -1.3 V for 10^{12} and 10^{13} cm^{-2}, respectively. For $N_n = 0$ cm^{-2}, $V_n = -3.4$ V would be obtained. Clearly the effect of the voltage drop across the Si-rich SiO_2 layer can be important with very thin SiO_2 layers and can influence the injection efficiency.

Figure 13 shows the I-V reproducibility histograms for the samples with injectors and the destructive breakdown histogram for the control sample with 207 Å of thermal oxide. In this case, both the field enhancement effect and voltage drop across the injector layer are important as discussed previously. The breakdown characteristics and the I-V reproducibility are improved by adding the Si-rich oxide layers. In this case, the best I-V reproducibility is obtained with $R_0 = 1$.

C. Breakdown and yield

The yield was defined as the percentage of the capacitors with a current of 5 μA (9.6×10^{-2} A/cm^2) at average electric fields larger than 2 MV/cm. Figures 14,

15, and 16 show the yield as a function of R_0 for different dual dielectric structures. This yield is for the initial (t=0) breakdowns. After the second voltage sweep, the yield of the control sample goes to approximately near zero due to the destructive breakdown of the sample. However, for the injector samples almost the same yield was obtained for the second measurement. For 200 Å Si-rich oxide/530 Å SiO_2 systems (Fig. 14), the best yield is obtained at $R_0=3$. On the other hand for 500 Å Si-rich oxide/530 Å SiO_2 systems (Fig. 15), the best yield is obtained at $R_0=10$. For thick (500 Å) Si-rich oxides with low excess Si content ($R_0=10$), the contribution of the traps (see the hysterisis in Fig. 2) in the Si-rich oxide to yield improvement seems to become significant. For 200 Å Si-rich oxide/thin SiO_2 (103 and 207 Å) systems (Fig. 16), the yield has a maximum at $R_0=2$. The reason for different values of R_0 at best yield and injection efficiency between composite structures with thin and thick gate SiO_2 layers seems to be due to the difference in structure or to the difficulty of controling the resistivity of the high excess Si content Si-rich oxide.

V. Conclusion

The yield, I-V reproducibility and injection efficiency of MOS structures using PECVD Si-rich oxide between the gate metal and SiO_2 layers were found to be improved compared to the control samples which have no Si-rich oxide layer. Optimum results were somewhat dependent on the composition of the Si-rich oxide. The optimum point varied somewhat for the different structures as shown in the Table 1 due to either variations in the field enhancement and/or voltage drop across the injector or due to the difficulty in controling the resistivity of the high excess Si content Si-rich oxide films. To decrease the voltage drop across the Si-rich oxide, the

thickness of Si-rich oxide should be thinner (100 Å or less). The high dielectric constant of the PECVD Si-rich oxide should be useful for application to dual-dielectric storage capacitors in dynamic memories (4).

ACKNOWLEDGMENT

REFERENCES

(a)   Present adress: Department of Electrical Engineering, Hiroshima University, Higashihiroshima 724, Japan

(b)   Present adress: INTEL CORPORATION, 3601 Juliette Ln., Santa Clara, CA 95051.

1.   D. J. DiMaria, K. M. DeMeyer, C. M. Serrano, and D. W. Dong, J. Appl. Phys. 52, 4825 (1981).

2.   *D. J. DiMaria, K. M. DeMeyer, and D. W. Dong, IEEE Trans. Electron Dev. ED-28, 1047 (1981).*

3.   D. J. DiMaria, D. W. Dong, C. Falcony, and S. D. Brorson, IEEE Electron. Dev. Lett. EDL-3, 191 (1982).

4.   S. K. Lai, D. J. DiMaria, and F. F. Fang, *to be published in IEEE Trans. Electron Dev.* (1983).

5.   D. J. DiMaria, D. W. Dong, C. Falcony, T. N. Theis, J. R. Kirtley, J. C. Tsang, D. R. Young, F. L. Pesavento, and S. D. Brorson, *to be published in J. Appl. Phys.* (1983).

6.   D. J. DiMaria and D. W. Dong, J. Appl. Phys. 51, 2722 (1980).

7.   D. J. DiMaria, R. Ghez, and D. W. Dong, J. Appl. Phys. 51, 4830 (1980).

8. A. M. Goodman, G. Harbeke, and E. F. Steigmeir, in Physics of Semiconductors 1978, ed. by B. L. H. Wilson (The Institute of Physics, Bristol and London, 1979), p. 805.

9. A. Hartstein, J. C. Tsang, D. J. DiMaria, and D. W. Dong, Appl. Phys. Lett. 36, 836 (1980).

10. E. A. Irene, N. J. Chou, D. W. Dong, and E. Tierney, Electrochem. Soc. 127, 2518 (1980).

11. J. M. Gibson, in Electron Microscopy 1980, ed. by P. Brederoo and G. Boom (The Seventh European Congress on Electron Microscopy Foundation, Leiden, 1980), pp. 310-311, Vol. 1.

12. D. J. DiMaria and D. R. Kerr, Appl. Phys. Lett. 27, 505 (1975).

13. H. S. Lee, Appl. Phys. Lett. 31, 475 (1977).

14. R. M. Anderson and D. R. Kerr, J. Appl. Phys. 48, 4834 (1977).

15. M. Koyanagi, Y. Sakai, M. Ishihara, M. Tazunoki, and N. Hashimoto, IEEE Trans. Electron Dev. ED-27, 1596 (1980).

16. K. Ohta, K. Yamada, K. Shimizu, and Y. Tarui, IEEE Trans. Electron Dev. ED-29, 368 (1982).

17. D. W. Dong, E. A. Irene, and D. R. Young, J. Electrochem. Soc. 125, 819 (1978).

18. Z. A. Weinberg, W. C. Johnson, and M. A. Lampert, J. Appl. Phys. 47, 248 (1976).

19. M. Lenzinger and E. H. Snow, J. Appl. Phys. 40, 278 (1969).

20. S. Yokoyama, N. Kajihara, M. Hirose, Y. Osaka, T. Yoshihara, and H. Abe, J. Appl. Phys. 51, 5470 (1980).

21. C. A. Neugebauer, J. F. Burgess, R. E. Joynson, and J.L. Mundy, J. Appl. Phys. 43, 5041 (1972).

Table I

Optimum gas phase ratios for injection efficiency, I-V

reproducibility, and yield for different dual-dielectric

composite structures.

SAMPLE		OPTIMUM GAS PHASE RATIO		
Si-Rich	SiO_2	INJ. EFFI.	I-V REP.	YIELD
200 $\overset{\circ}{A}$	530 $\overset{\circ}{A}$	1	3	3
500 $\overset{\circ}{A}$	530 $\overset{\circ}{A}$	1	3	10
200 $\overset{\circ}{A}$	207 $\overset{\circ}{A}$	2	1	2
200 $\overset{\circ}{A}$	103 $\overset{\circ}{A}$	2	1	2
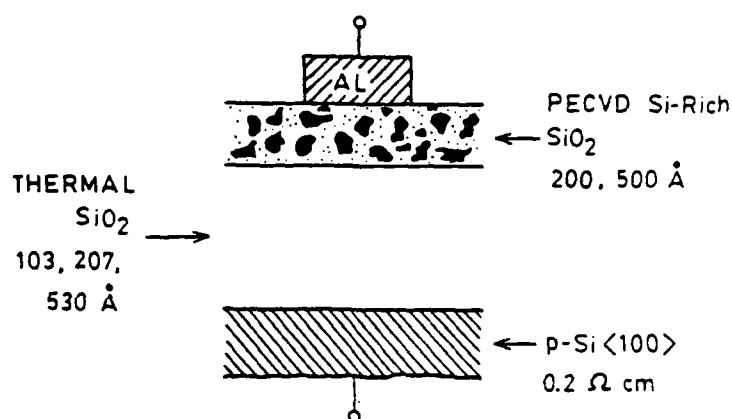
Fig. 1   Schematic illustration of a capacitor structure with a dual dielectric composite incorporating a Si-rich SiO_2 layer on top of thermal SiO_2.
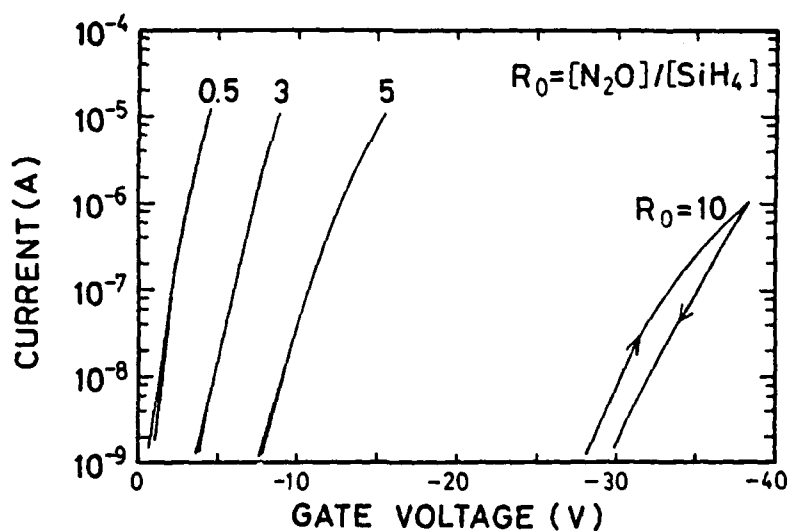
Fig. 2   Magnitude of the dark current at room temperature as a function of the ramped gate voltage ($|dV_g/dt| = 0.5$ V/sec) for Si-rich oxides with a thickness of 435 Å of various compositions from $R_0 = 0.5$ to 10 inserted between the Al gate and the p-type single crystal silicon substrate. The gate area is 5.19×10^{-3} cm^2. These samples were annealed in N_2 at 1000°C for 30 min prior to gate electrode deposition.
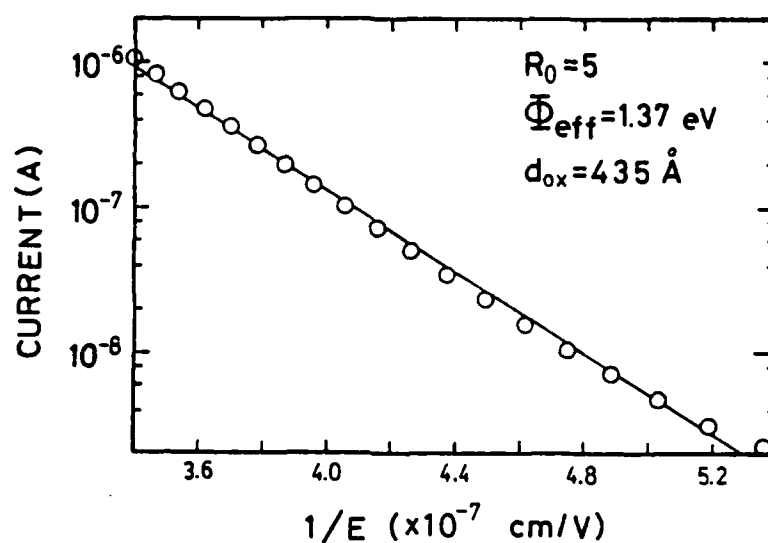
Fig. 3   *Magnitude of the current density at room temperature as a function of one over the magnitude of the average electric field for negative gate voltage for the Al/Si-rich oxide ($R_0=2$, 435 Å in thickness)/p-Si structure similar to those in Fig. 2. The full line represents a least-squares fit to all plotted lnI vs. 1/E data. If the measured currents were due to Fowler-Nordheim tunneling from a planar interface, an effective barrier height of $\Phi_{eff}=1.37$ eV is deduced from this data using a least-squares fit with a tunneling electron mass$=0.5 \times m_0$ where m_0 is the free electron mass.*
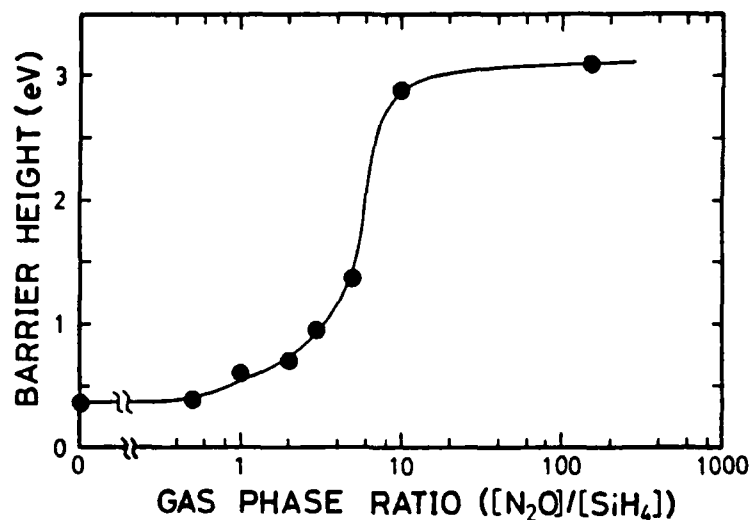
Fig. 4   Effective barrier heights for Al/Si-rich oxide/p-Si structures as a function of R_0 deduced from least-squares fitting of the I-E data as described in Fig. 3.
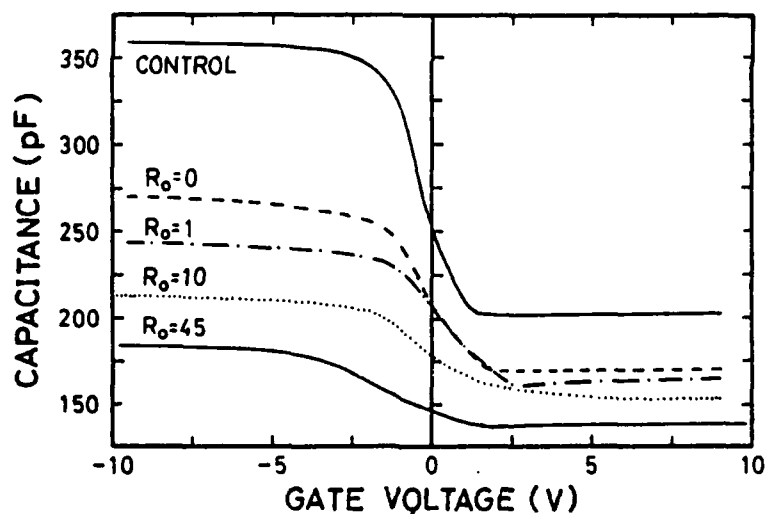
Fig. 5   1 MHz, small signal (15 mV), capacitance-voltage curves for the Al/Si-rich oxide (200 Å in thickness)/thermal SiO_2 (530 Å in thickness)/p-Si structures with different R_0 as compared to a control sample with a 530 Å thick SiO_2 layer and no Si-rich oxide layer. These samples were annealed in N_2 at 1000°C for 30 min prior to gate electrode deposition.
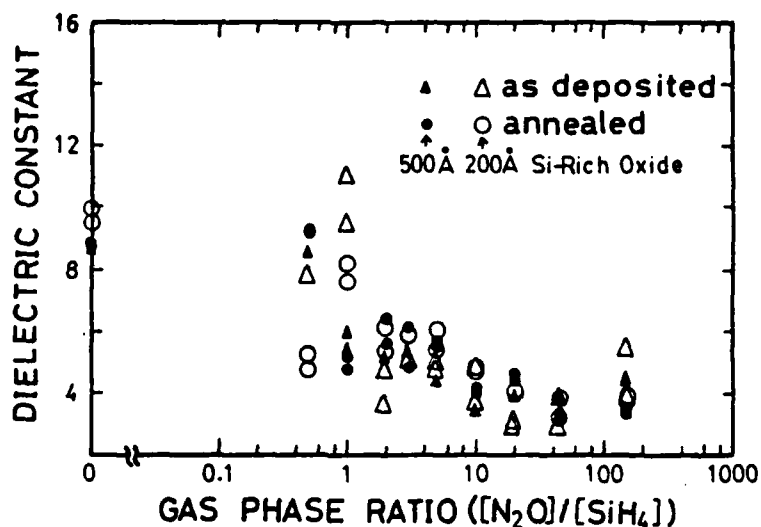
Fig. 6   Dielectric constants of the Si-rich oxide films (200 Å and 500 Å in thickness) on top of a 530 Å thermal oxide layer deduced from 1 MHz capacitance-voltage data similar to that in Fig. 5.
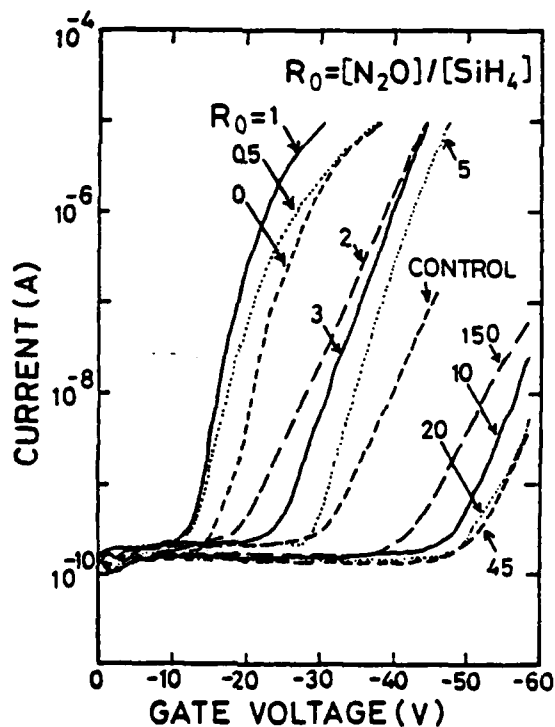
Fig. 7   Typical ramped I-V curves for Al/Si-rich oxide (200 Å in thickness)/SiO$_2$ (530 Å in thickness)/p-Si structures (see Fig. 1) for various injectors with different R$_0$. The control structure is Al/SiO$_2$ (530 Å in thickness)/p-Si. These samples were annealed at 1000°C in N$_2$ for 30 min prior to gate electrode deposition.
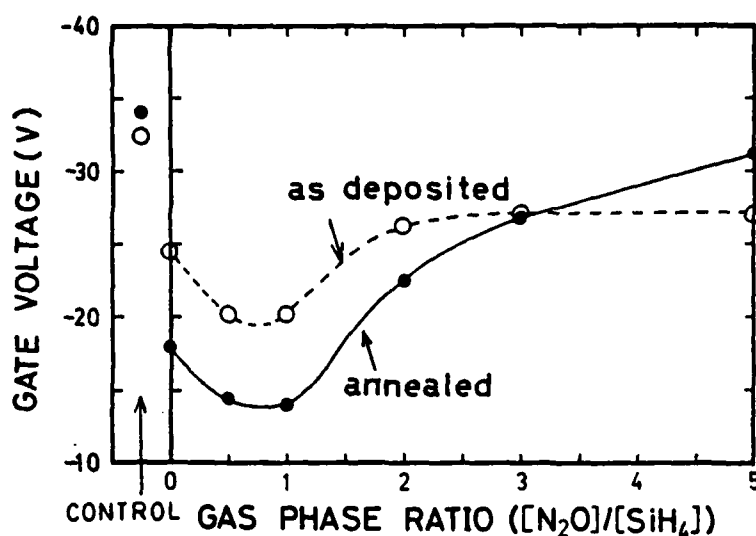
Fig. 8   Gate voltages at I=1 nA (1.92×10⁻⁷ A/cm²) for the same type structures as in Fig. 7 with and without high temperature annealing (1000°C - N₂ - 30 min. prior to gate electrode deposition).
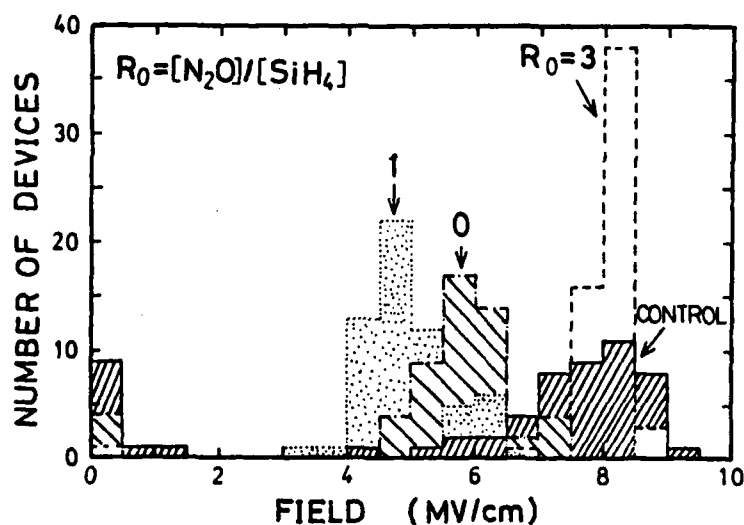
Fig. 9   Histograms in 0.5 MV/cm bins of the number of capacitors (5.19×10⁻³ cm² in area) on Al/Si-rich oxide (200 Å in thickness)/SiO₂ (530 Å in thickness)/p-Si composite structures (see Fig. 1) with different R_0 and on a control wafer (no Si-rich oxide injector) with a current of 5×10⁻⁶ A (9.6×10⁻⁴ A/cm²) as a function of the magnitude of the average electric field in the SiO₂ layer. Samples were ramped with negative gate voltage from 0 V at a ramp rate magnitude of 1 MV/cm-sec. These samples were annealed in N₂ at
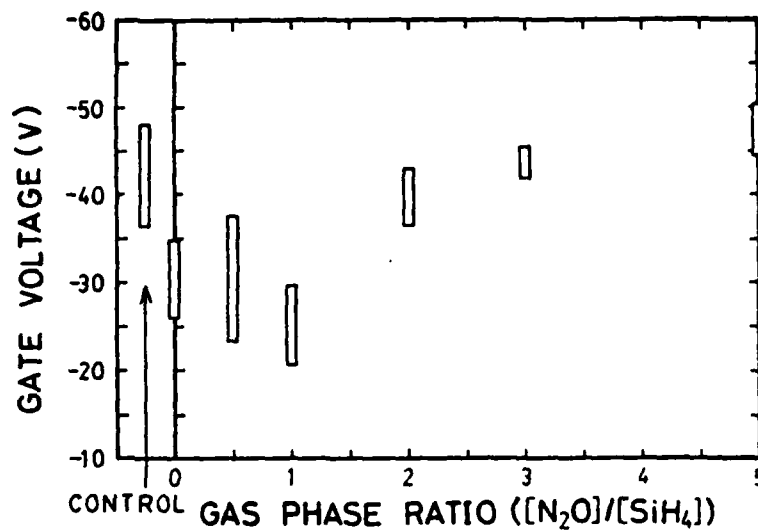
Fig. 10   Gate voltages at the peak of the histograms shown in Fig. 9 as a function of R_0. The error bar indicates the half width of the distribution for the histogram.
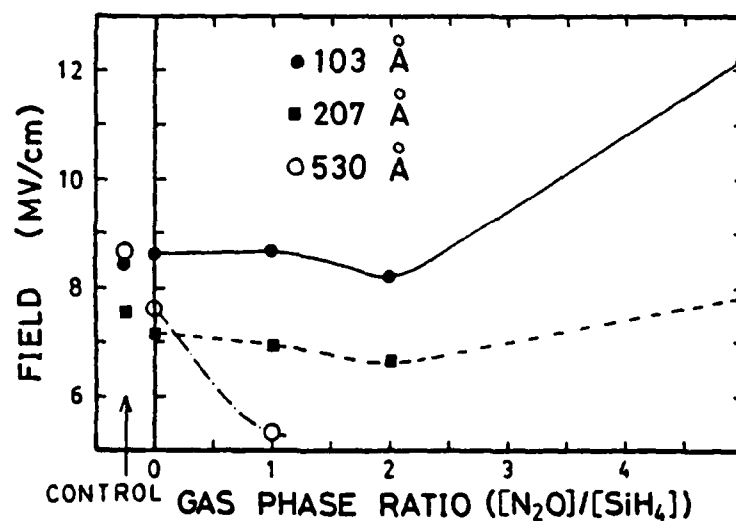

Fig. 11   The average SiO_2 electric field at 1 nA (4.83×10^{-7} A/cm²) as a function of R_0 for 103, 207, and 530 Å oxide samples with Si-rich oxide injectors of 200 Å thickness (see Fig. 1) annealed at 1000° in N_2 for 30 min. prior to gate electrode deposition.
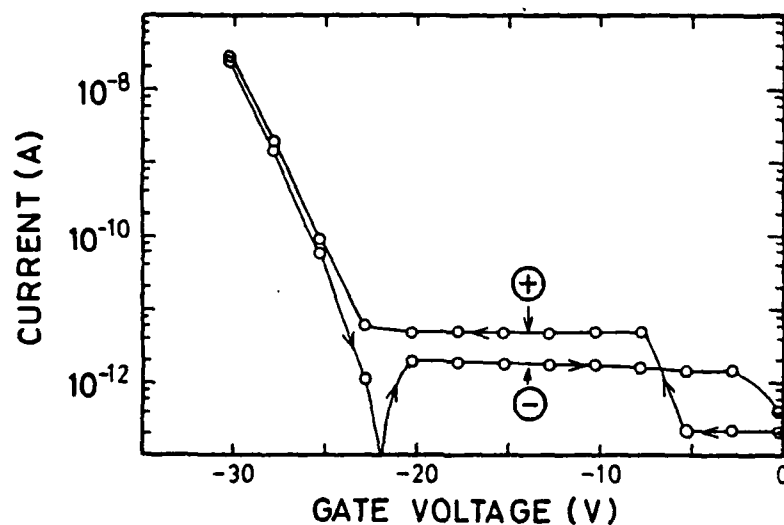
Fig. 12   Point-by-point magnitude of the dark current as a function of negative

gate voltage on a Al/Si-rich oxide ($R_0 = 1$, 200 Å in thickness)/SiO$_2$

(530 Å in thickness)/p-Si structure (see Fig. 1) which was annealed at

1000°C in N$_2$ for 30 min. prior to gate electrode deposition. In this

measurement the gate voltage was stepped in -2.5 V increments starting

from 0 V every 20 sec with the dark current being measured 18 sec

after each voltage step. The sign of the current flow is is indicated in

this figure. The gate area is 9.23×10^{-3} cm^2.
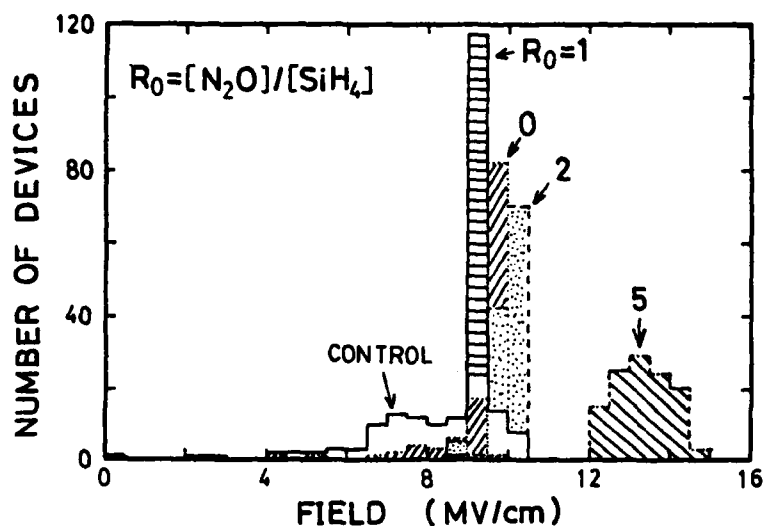
Fig. 13   Histograms in 0.5 MV/cm bins of the number of capacitors

$(4.13 \times 10^{-4}$ cm^2 in area) on Al/Si-rich oxide (200 Å in

thickness)/SiO$_2$ (207 Å in thickness)/p-Si composite structures (see

Fig. 1) with different R$_0$ and on a control wafer (no Si-rich oxide

injector) with a current of 5 μA $(1.21 \times 10^{-2}$ A/cm^2) as a function of

the magnitude of the average electric field in the SiO$_2$ layer. Samples

were ramped with negative gate voltage from 0 V at a ramp rate magni-

tude of 1.5 MV/cm-sec. These samples were annealed at 1000°C in

N$_2$ for 30 min. prior to gate electrode deposition.
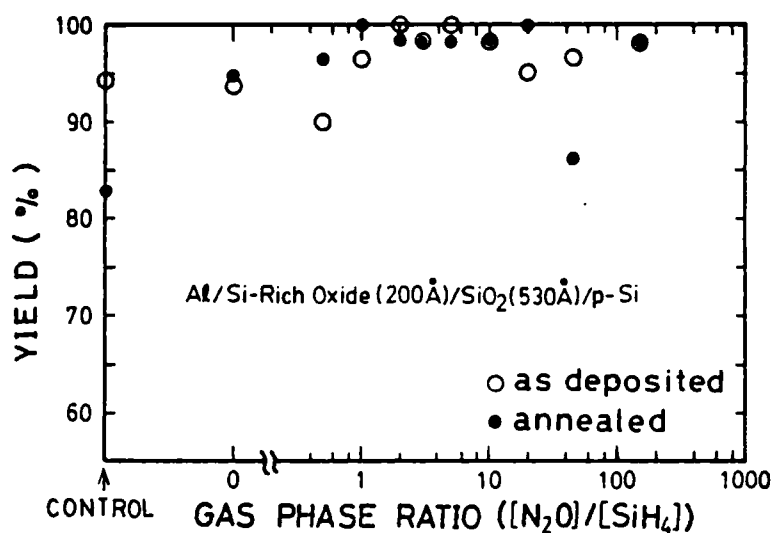
Fig. 14   The yield for Al/Si-rich oxide (200 Å in thickness)/SiO₂ (530 Å in
           thickness)/p-Si structures (see Fig. 1) with and without high tempera-
           ture annealing (1000°C - N₂ - 30 min. prior to gate electrode depos-
           ition) as a function of R_0. The yield is defined in the text.
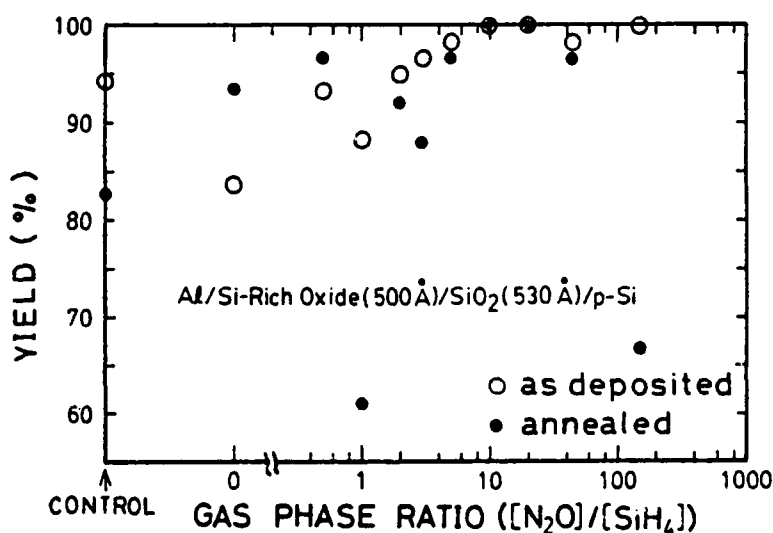

Fig. 15   The yield for Al/Si-rich oxide (500 Å in thickness)/ᶜiO₂ (530 Å in
           thickness)/p-Si structures (see Fig. 1) with and without high tempera-
           ture annealing (1000°C - N₂ - 30 min. prior to gate electrode depos-
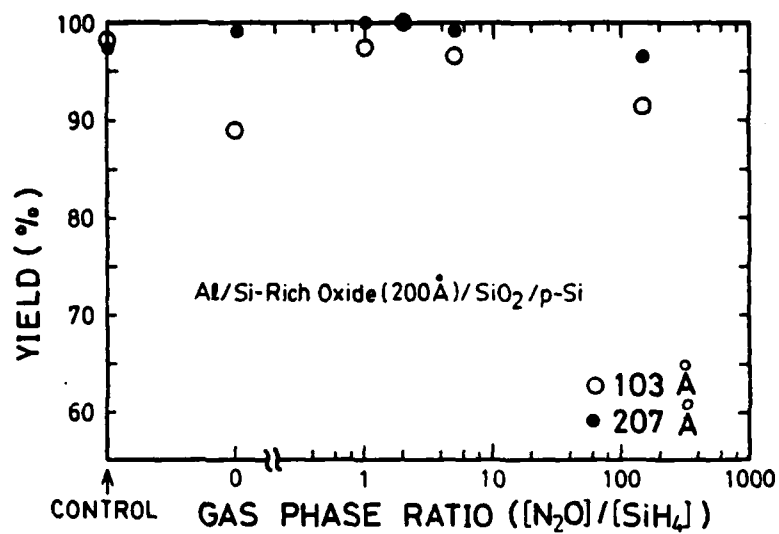   ition) as a function of R_0.

Fig. 16   The yield for Al/Si-rich oxide (200 Å in thickness)/SiO$_2$ (103 and 207

Å in thickness)/p-Si structures (see Fig. 1) with high temperature

annealing (1000°C - N$_2$ - 30 min. prior to gate electrode deposition) as

a function of R$_0$.

Charge Transport and Trapping Phenomena in Off-Stoichiometric SiO$_2$ Films*

D.J. DiMaria, D.W. Dong, C. Falcony[a], T.N. Theis, J.R. Kirtley, J.C. Tsang, D.R. Young and
F.L. Pesavento
I.B.M. Thomas J. Watson Research Center
Yorktown Heights, New York 10598

and S.D. Brorson
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Typed by Stella B. Havreluk and Maryann Pulice

Abstract: The electrical characteristics of off-stoichiometric SiO$_2$ films have been investigated. The off-stoichiometric oxide films studied had an excess atomic Si content in the range of 1% to 6%. Raman spectroscopy and photoconductivity measurements indicate that the excess Si is present as amorphous Si islands or small crystallites embedded in SiO$_2$ forming a two-phase material. These films differ in structure from previously reported films where dual dielectric layers of stoichiometric SiO$_2$ and Si-rich SiO$_2$ with $\geq$ 13% excess Si were used. These dual dielectric films were observed to produce electron injection from contacting electrodes via the Si-rich SiO$_2$ layer into the SiO$_2$ layer at lower average electric fields. This injection mechanism was believed to be due to localized electric field enhancement near the SiO$_2$—Si-rich SiO$_2$ interface caused by the curvature of the tiny Si islands in the SiO$_2$ matrix. The current vs. voltage characteristics of the off-stoichiometric oxide films which will be discussed here were found to be highly non-ohmic, showing an increase in conductivity with increasing excess silicon content in the material. At low fields ($\leq$ 1 MV/cm), these films have a very small conductance with leakage current densities smaller than 10^{-11} A/cm^2 at room temperature. Furthermore, the effect of permanent charge trapping was observed to decrease in these films with increasing Si content. It is proposed that the electron transport across the film is controlled primarily by tunneling between the silicon islands with a reversible space charge region $\leq$ 150 Å in extent near the injecting contact, limiting the current measured in the external circuit. It is also proposed that the decrease in permanent electron trapping is associated with this conduction mechanism and/or the possibility of trapped electrons in the SiO$_2$ phase tunneling to the Si islands. With the incorporation of these off-stoichiometric oxides into electrically-alterable read-only-memory devices, extended write/erase cycling (by at least four orders of magnitude) beyond that normally observed for equivalent devices using stoichiometric SiO$_2$ layers is demonstrated due to the reduction in permanent electron trapping in the oxide layer and its effect on the electric fields.

I. Introduction

Permanent charge build-up in SiO_2 layers represents a limitation for the long term performance and reliability of electronic devices that depend upon charge transfer across thin oxide films for their operation. This is the case for non-volatile memories such as the electrically-alterable read-only-memory (EAROM) structures in which charge is transferred to and from a floating gate through a thin oxide layer.[1] Capture of electrons into energetically deep traps in the SiO_2 is responsible for the permanent nature of this charge build-up in the oxide.[2-7] Extensive studies in SiO_2 films have shown that these traps are associated with the presence of water related impurities[4,5] and/or possibly with non-mobile sodium[6,7] in the oxide layer. The electron capture efficiency in thermally grown and chemically-vapor-deposited (CVD) oxide layers incorporated in MOS structures is greatly reduced with post-oxidation and post-metallization thermal annealing treatments.[6,8,9] Similar effects have been demonstrated with the annealing at 1000°C in N_2 of the intervening oxide layer in a dual electron injector structure (DEIS) in EAROM devices.[1] However, the basic problem of permanent capture of electrons in the oxide layer has remained unsolved. An alternative approach to solve this problem has been to reduce the oxide layer thickness ($\leq$ 70 Å but $\geq$ 40 Å) to allow the trapped charge to tunnel out to the electrical contacts under the influence of its own internal electric field and/or the applied electric field.[10] This approach has the disadvantage that the trapped charge relaxation process is slow. Furthermore, the difficulties associated with the preparation of uniform oxide films in this thickness range makes this solution cumbersome. Slightly off-stoichiometric oxides (1% to 6% excess silicon) which show a very pronounced decrease in the permanent trapped charge build-up problem will be discussed in detail in this article, and they will be shown to offer a better solution to reduce permanent trapped charge build-up in the oxide layer of devices than any of the above mentioned procedures (thin SiO_2 or annealing).

Sections II and III will describe different optical and electrical measurements including photoconductivity, dark conductivity, capacitance, electroluminescence, and carrier separation to investigate the physical mechanisms involved in producing the observed enhanced conductivity and low level of permanent trapping of the off-stoichiometric oxides containing $\leq$6% excess atomic Si. Material characterization tools such as Raman spectroscopy to investigate the two phase nature (Si and SiO_2) of these materials and scanning electron microscopy (SEM) to investigate surface roughness and uniformity will also be discussed in these sections. Experiments on capacitor structures with and without electron injector layers[1,11-15] will also be described. Electron injector layers are thin layers of Si-rich SiO_2 ($\geq$13% excess atomic Si) which, when placed in between a stoichiometric SiO_2 layer an a contacting electrode, give

93

field-enhanced electron injection from the irregularly-shaped Si islands into the oxide layer at moderate average electric fields without destructive breakdown.[1,13-15] The dependence of dark currents or photocurrents on temperature, contacting electrode material, and oxide thickness will be presented in these sections. In section IV, all the experimental observations will be united into a plausible physical picture which involves dominant electron conduction by direct tunneling between Si islands in the bulk of off-stoichiometric oxides with reversible space charges near the contact or injector interfaces at least partially screening the electric fields. The lack of permanent charge trapping will be related to the conduction mechanism with few electrons getting into the SiO_2 conduction band and/or the possibility of direct tunneling of energetically-deep, trapped electrons to nearby Si islands. Finally the use of these materials in EAROM devices to increase cyclability and to reduce operating voltages will be briefly covered in section V as an example of the relevance of mixed phase oxide materials in the electronics industry.

II. Experimental

A) Sample Preparation

The off-stoichiometric oxide layers were deposited using chemical vapor deposition (CVD) techniques.[11-13] The excess silicon is introduced by adjusting the ratio (R_o) of the concentration of N_2O to SiH_4 in the gas phase. A 1% to 6% excess silicon content in the oxide was obtained by setting the R_o value in the range of 50 to 30, as deduced from Rutherford backscattering (RBS).[16] A stoichiometric SiO_2 film requires $R_o \geq 100$ while Si-rich SiO_2 injectors with an excess atomic Si content of 13% to 15% requires $R_o = 3$.[13] Film thickness and refractive index were deduced from ellipsometry measurements using techniques previously described.[11] The samples used for ramp I-V, the point-by-point I-V, flatband voltage, avalanche injection, electroluminescence, photoconductivity, and photodetrapping measurements were MOS type capacitors involving layers of stoichiometric or off-stoichiometric oxides, dual-electron-injector-structures (DEISs), or single-electron-injector-structures (SEISs) deposited on <100> n-type or p-type silicon substrates of various resistivities (2, .2, .001 Ωcm) appropriate for the measurement. The different types of capacitor configurations are shown in Fig. 1. Figure 2 shows the two types of device configurations used in this study. The DEIS EAROM device in Fig. 2a will be discussed in section V, and the off-stoichiometric oxide field-effect-transistor (FET) in Fig. 2b will be discussed in section III-F. This structure (Fig. 2b) was used to separate hole and electron contributions to the current. The circular metal electrodes were thermally evaporated aluminum or gold layers with an area of .005 cm^2

or .01 cm^2 (for electroluminescence). The electrodes for those samples used in photodetrapping, photoconductivity, and electroluminescence measurements were 135 Å to 250 Å thick in the case of Al and 250 Å to 400 Å thick for Au electrodes; the rest of the samples had electrode layers thicker than 1000 Å. The structures used for photodetrapping measurements (Fig. 1b) consisted of an off-stoichiometric layer (R_o=50 or R_o=10) sandwiched between two stoichiometric SiO_2 layers (each 400 Å thick), where the bottom oxide layer was thermally grown on the silicon substrate while the top layer was CVD oxide. The SEIS (Fig. 1c) or DEIS (Fig. 1d) stacks consisted of one or two Si-rich SiO_2 injectors, respectively, with $\geq$ 13% excess atomic Si with an oxide layer which was either stoichiometric or off-stoichiometric. All samples were annealed at 1000°C in an N_2 ambient for 30 min. before metallization, unless otherwise specified. A post-metallization anneal at 400°C for 20 min. in forming gas was given to all samples except those with thin gold electrodes. The Raman spectra measurements were performed on 2 μm thick off-stoichiometric oxides (R_o=30) deposited on a sapphire substrate with no metal counter electrode present.

The EAROM devices (Fig. 2a) were fabricated using a self-aligned, double polycrystalline silicon (poly-Si) gate process on 0.5 Ω-cm <100> p-type Si wafers. The details of the device fabrication process have been previously described.[14,15] The device data presented here were taken on 3-port devices with a floating gate area A_2 = 2.5 x 10^{-6} cm^2 and a control gate area A_1 = 1.3 x 10^{-6} cm^2 (at the masking level) of n-degeneratively doped poly-Si. The gate oxide thickness was 650 Å and 850 Å for devices XDEISII 3-B,G-16 and XDEISII 4-B,E,G,I-16, respectively. The intervening oxide layers in the DEIS or modified DEIS (MDEIS) stack were either 300 Å or 600 Å with varying amounts of excess Si; wafers XDEISII 3-G and XDEISII 4-B had intervening stoichiometric SiO_2 layers with R_o = 200 (0% excess atomic Si), XDEISII 3-B and XDEISII 4-E had oxide layers with R_o = 50, (1-2% excess atomic Si), XDEISII 4-G had an oxide layer with R_o = 40 (3-4% excess atomic Si), and XDEISII 4-I had an oxide layer with R_o = 30 (5-6% excess atomic Si). The silicon rich injectors were 300 Å and 200 Å thick for devices from the XDEISII 3 and XDEISII 4 series of wafers, respectively. All DEIS or MDEIS stacks were annealed in N_2 for 30 min. at 1000°C after deposition.

The FETs used for the carrier separation measurements used a self-aligned, single polysilicon gate process on 0.5 Ωcm <100> p-type Si wafers, similar to that used for the double polysilicon gate EAROM devices described previously. The n-degenerately doped poly-Si control gate area of these large devices (MDT-HOTEL 1-A,B,E,F,G,H-42) was 1 x 10^{-4} cm^2 at the masking level. The gate insulators were formed from either CVD off-

stoichiometric $R_o = 50$ (1-2% excess atomic Si) oxides with thicknesses of 300 Å or 600 Å for wafers MDT-HOTEL 1-G and H, or from the same CVD $R_o = 50$ layers as on G and H but with a 70 Å thick thermally-grown (at 1000°C in O_2 with 4.5% HCl) stoichiometric SiO_2 layer between the Si substrate and the thicker off-stoichiometric oxide layer for wafers MDT-HOTEL 1-A and B, or from a 245 Å or 670 Å thick thermally grown (at 1000°C in O_2 with 4.5% HCl) SiO_2 layer only for wafers MDT-HOTEL 1-E and F. All gate oxides were annealed in N_2 for 30 min. at 1000°C after CVD deposition or SiO_2 growth. These devices also had an Al guard ring over the field oxide region which was usually kept grounded.

B. Measurement Techniques

The photocurrent and photodetrapping measurements[13,17] were performed using a 900 W Xenon lamp that provided a continuous spectrum in the photon energy range from 2 eV to 5.5 eV. The light was passed through a Bausch and Lomb 500-mm grating monochromator with a reciprocal linear dispersion of 3.3 mμ/mm set for a 5-mμ bandpass before being focused on the sample, and cut-off filters were used to remove second and higher order light. The photocurrents induced in the MOS structures (Fig. 1a) as a function of photon energy were measured with a fast picoammeter (Keithley Model 417) for different applied voltages. The photodetrapping measurements involved an initial electron injection by internal photoemission[2] into the Si/SiO_2/R_o=50 or 10 off-stoichiometric SiO_2/SiO_2/metal structures (see Fig. 1b) to fill the off-stoichiometric oxide with charge. The photoinjection current was induced from the metal contact (negative polarity on the gate) using light with photon energy of 4.1 eV. The amount of charge trapped in the off-stoichiometric oxide was obtained by measuring the flatband voltage shifts of a high frequency (1 MHz), small signal (15 mV) capacitance vs. voltage curves. Subsequently, a selective photodetrapping of the trapped charge as a function of photon energy was measured with a point-by-point sequence in which the light of a given photon energy was focussed on the sample for 1 min. and then the change in the flatband voltage was measured. A Boonton capacitance-inductance meter Model 71 A-SI and a voltage tracking circuit were used to monitor the shifts of the flatband voltage.[5,8,9] The capacitor area (.005 cm^2) was always uniformly illuminated by the larger focused light spot.

The ramp I-V measurements were performed by applying a voltage ramped at a constant rate ($|dV_g/dt| \approx 0.5$ V/sec) to the sample. The current through the devices was measured with a log-picoammeter (Keithley Model 26000). The temperature dependence of the ramp I-V characteristics was measured using a cryostat coupled with a liquid-nitrogen

transfer system (Air-products Heli-tran) for temperatures in the range of 80°K to room temperature (292°K). A resistance heated stage was used for temperatures above 22°C up to 300°C. For current measurements where the voltage polarity could drive the Si substrate into deep depletion, in particular p-Si substrates at large positive gate voltages, the samples were illuminated with white light. This insures rapid formation of an inversion layer at the substrate Si-oxide interface, negligible voltage drop across the Si layer, and an adequate supply of carriers at this interface which can tunnel into the oxide. The white light generates minority carriers (electrons for p-Si) around the periphery of the opaque gate electrode which flow laterally into the region under the electrode. However, this white light illumination had little effect on the conductivity of the off-stoichiometric oxides themselves over the electric field range of interest.

For studies of the time dependence of the dark currents, a single voltage step was applied to the gate of MOS-type structures, and the current was measured using a Keithley 427 fast current amplifier and then displayed on a Tektronix 7623A storage oscilloscope. The voltage polarity used was negative for p-Si substrates and positive for n-Si substrates so that the Si was always in accumulation during the measurement.

Avalanche injection experiments were performed at room temperature by driving the Si substrate into depletion using sawtooth-shaped voltage pulses until the Si broke-down, thereby producing some "hot" electrons by avalanche multiplication in the silicon bulk which could energetically surmount the Si-oxide interfacial energy barrier.[8] The sawtooth-shape for the voltage pulses was used to minimize high field stressing of the oxide layer after the production of the avalanche plasma in the Si and the subsequent collapse of the voltage dropped across the Si depletion region. Periodically, the avalanche current into the oxide layer was shut off and the flatband voltage was measured automatically. The instrumentation for these experiments has been described previously.[8]

The luminescence experiments were performed at room temperature using an Instruments for Scientific Analysis (ISA) single-pass monochromator (147 lines/mm grating) with the entrance and exit slits set at 1 mm. Mirror optics were used to focus the light emitted from the oxide capcitors or SEISs depicted in Figs. 1a and 1c, respectively, on the monochromator entrance slit and an RCA No. 31034 photon tube with photon-counting was used at the monochromator exit slit.

The Raman spectra were excited by the 5145 Å line of an Ar^+ laser at room temperature and measured using backscattering geometry. The scattered light was analyzed by a conventional double monochromator and detected using photon counting techniques.

The ellipsometer used to determine film thickness and refractive index was a Roldoph Research manual ellipsometer with the optical elements arranged in a polarizer-compensator-analyzer configuration. Two zone measurements were made on ~ 1000 Å thick films deposited on single crystal Si substrates using the 5461 Å wavelength light from a Hg lamp.

The experimental arrangement for the EAROM FET measurements has been described in detail earlier.[1] The drain to source current was measured with a fast current amplifier (Keithley Model 127) while a ramp voltage was applied to the gate with the drain baised at +0.1 V. Both gate voltage and drain to source current were displayed on a storage scope. The measurement was performed in fractions of a second to minimize read-perturb effects. The write/erase operation of the devices was carried out using square voltage pulses from two Hewlett-Packard pulse generators (Model 214A).

III. Experimental Results

A. Two Phase Characteristics and Photoconductivity

Raman spectra from annealed ($1000°C$ in N_2 for 30 min) and unannealed off-stoichiometric oxides ($R_o = 30$) are illustrated in Fig. 3. Both samples show the broad Raman band near 500 cm^{-1} which is characteristic of Raman scattering from amorphous silicon clusters in SiO_2.[18,19] The Raman lines due to the underlying sapphire substrate have been suppressed in the figure. This result is in contrast with previous work on Si-rich oxides ($R_o = 3$) with 13% to 15% excess silicon in which the high temperature ($\geq 1000°C$) annealing of the samples resulted in Raman spectra with a component similar to that from single crystal or polycrystalline silicon.[18,19] From the strength of the Raman scattering signal and the fact that no observable difference exists between the annealed and unannealed samples, it is suggested that the excess silicon groups in either microscopic crystalline clusters $\leq$ 30 Å in diameter or microscopic amorphous regions,[20] or both. Recent work by Veprek and co-workers using x-ray diffraction and Raman scattering techniques[21,22] suggests that there is an abrupt change from a crystalline silicon phase to and amorphous siliconqphase for Si crystallites smaller than 30 Å in stress free films and smaller than 20 Å in films under compressive stress. Recent preliminary x-ray photoelectron spectroscopy (XPS) measurements also show

that there are at least two distinct phases (Si and SiO_2) in unannealed Si-rich SiO_2 films with excess atomic Si from 13-15% (R_o = 3) down to 5-6% (R_o = 30).[23]

Transmission electron microscopy (TEM) on 500Å to 1000Å thick, CVD annealed (1000°C in N_2 for 15 to 30 min.) R_o = 30 to 50 (5-6% to 1-2% excess atomic Si) oxide films showed no observable Si particles.[24] These samples were prepared by depositing the films on planar Si substrates and then etching away a portion of the Si substrate, leaving a free standing film of the off-stoichiometric oxide. Previous TEM measurements on annealed CVD R_o = 3 ($\geq$13% excess atomic Si) material revealed Si crystallities in these films, with the largest crystallities having a diameter of $\approx$ 50 Å.[25,26] The films were annealed at 1000°C to convert some of the amorphous Si island regions into Si crystallities which could then be observed using TEM.[18,19,25,26] Since no crystallities were observed on the R_o = 30 to 50 films, it appears that the crystallities are smaller than 50 Å, and/or the annealing has not converted the amorphous Si islands to crystallities. However, the former possibility (Si islands smaller than 50 Å in diameter) seems more likely from various experimental data that will be presented and discussed in this article.

The presence of the excess silicon in a second phase in the off-stoichiometric SiO_2 is supported by the results obtained from photoconductivity and photodetrapping experiments on this material. The photoresponse Y is obtained from the measured photocurrent I_{photo} using the expression

$$Y = \frac{I_{photo}\, \hbar\omega}{SA} \quad \text{(photoelectrons/photon)} \qquad (1)$$

where $\hbar\omega$ is the incident photon energy, A is the device area, and S is the average light intensity (power per unit area) in the bulk of the oxide layer corrected for optical interference effects in the multilayer structure and for absorption in the metal electrode and Si substrate.[27] Figure 4 shows the cube root of the photoresponse vs. the incident photon energy for an MOS structure incorporating a 1200 Å thick off-stoichiometric oxide (R_o=50) and a 135 Å thick semitransparent Al electrode (see Fig. 1a) The different curves correspond to photocurrent measurements obtained by applying $\pm$9 V or $\pm$12 V to the gate. The photoresponse behavior with photon energy was observed to be relatively independent of the applied voltage polarity (as shown in Fig. 4), of the oxide thickness in the range from 300 Å to 1200 Å, or of the gate metal (Al or Au as seen by comparing Figs. 4 and 5) indicating a bulk rather than interface limited phenomena. The photoresponse shows two thresholds with photon energy at $\approx$ 3 eV and $\approx$ 4 eV. This spectral response is similar to that observed from internal photoemission of electrons from single crystal silicon into SiO_2.[13,28,29] Therefore, these results suggest that the

photocurrents observed are generated by optical excitation of electrons from near the bottom of the conduction band (for $\approx$ 3 eV threshold) or near the top of the valence band (for $\approx$ 4 eV threshold) of the silicon islands to near the bottom of the conduction band of the SiO_2. Figure 5 shows the temperature dependence of the chopped light a-c photoconductivity on a 1200 Å R_o=50 off-stoichiometric oxide film incorporated into an MOS capacitor with a thin Au (400 Å thick) electrode under a voltage bias of -20 V. Again the 3 eV and 4 eV thresholds are seen. These thresholds do not appear to be strongly temperature dependent as would be expected for excitation from potential wells which are deep with respect to thermal energy kT ($\sim$ .05 eV for 300°C), where k is Boltzmann's constant (8.62 $\times$ 10^{-5} eV/°K) and T is the absolute temperature in °K.

The photoionization measurements of charge trapped on the silicon islands (previously charged by internal photoemission) in an off-stoichiometric oxide sandwiched between stoichiometric SiO_2 (see Fig. 2b) were performed by grounding both electrodes and illuminating the sample with monochromatic light in the photon energy range from 3 eV to 5 eV. Photo-excited electrons are swept out towards the electrical contacts under the influence of the internal electric field of the trapped electrons. The amount of charge detrapped at a given wavelength is measured by tracking the corresponding shifts of the flatband voltage V_{FB}. The shift in flatband voltage can be described using first order kinetics analysis by the expression[17]

$$\Delta V_{FB}(t) = \Delta V_{FB}(0) \exp\left[-\chi(\hbar\omega)\Sigma(\hbar\omega)t\right] \qquad (2)$$

where $\chi(\hbar\omega)$ is the convolution of the photon flux with the spatial distribution of the trapping centers, $\Sigma(\hbar\omega)$ is the effective photoionization cross section of the trapped charge (i.e., the convolution of the photoionization cross section with the optically accessible trap distribution over energy), and t is the illumination time. Figure 6 shows the cubic root of the photoionization cross section as a function of photon energy. The charge trapped in the intermediate off-stoichiometric SiO_2 layer (R_o=50, or R_o=10) is treated in this case as a trapping distribution in the stoichiometric oxide. The experimental data is compatible with the idea that the silicon islands behave as potential energy wells in the SiO_2 with optical activation energies of $\approx$ 3 eV and $\approx$ 4 eV. Notice the similarity of the data in Figs. 4-6.

Although some electric field dependence for either the photoconductivity or photodetrapping measurements was observed, the magnitude of the effect was relatively weak as seen by comparing the data in Figs. 4-6. This weak field dependence is consistent with either a Schottky barrier lowering effect and/or an internal Schottky (Poole-Frenkel) effect where the applied electric field and either the attractive potential of the image charge in a contacting

electrode (Schottky) or the attractive potential of an ionized site (Poole-Frenkel) creates a reduced energy barrier for an escaping charge carrier.[30] The average electric field or local field range explored in the photoconductivity or photodetrapping measurements, respectively, was between $\sim 2 \times 10^5$ and $\sim 2 \times 10^6$ V/cm. The average electric field for the photoconductivity measurements is defined as $E = (V_g - \phi_{ms} - \Psi_s)/\ell_o$ where V_g is the voltage applied to the gate electrode, ϕ_{ms} is the difference in work function between the gate material and the Si substrate (by convention, expressed in volts), Ψ_s is the surface potential for the substrate Si-oxide interface, and ℓ_o is the oxide thickness. The local electric field for the photodetrapping (photoionization) measurements where the gate and substrate are grounded was deduced from measurements of shifts of the flatband voltage produced by the trapped charge in the intervening off-stoichiometric oxide layer.[3,17] For SEIS or DEIS structures, the average electric field is still approximately valid since very little of the applied voltage is dropped across the Si-rich SiO_2 injectors which are very conductive (except at very low electric fields) with respect to the oxide layer. Over the electric field range from 2×10^5 to 2×10^6 V/cm, the predicted energy barrier lowering $\Delta\Phi$ was calculated using $\Delta\Phi_{PF} = q (qE/\pi \epsilon_i)^{1/2}$ for Poole Frenkel or $\Delta\Phi_S = q(qE/4\pi \epsilon_i)^{1/2}$ for Schottky[30] where q is the magnitude of charge on an electron (1.6×10^{-19} coul) and ϵ_i is the high frequency dielectric permittivity (assumed to be similar to SiO_2 where $\epsilon_i = (2.15) \times 8.86 \times 10^{-14}$ F/cm as discussed in Section III-C). These calculated values for $\Delta\Phi$ were found to be between $\sim.2$ eV and .7 eV for Poole-Frenkel and between $\sim.1$ eV and .35 eV for Schottky, respectively, over this electric field range. By extrapolating the quasi-linear portions of data such as those in Figs. 4-6 to a null value of $(\text{photoresponse})^{1/3}$ or $(\text{effective photoionization cross section})^{1/3}$, barrier energies for the range of samples and electric fields discussed here were deduced. $\Delta(\Delta\Phi)$ values [where $\Delta(\Delta\Phi) = \Delta\Phi(E_2) - \Delta\Phi(E_1) = \Phi(E_1) - \Phi(E_2)$ and the subscripts refer to different values of the electric field] were found in the range of $\leq.1$ to .3 eV from these data, consistent with the predicted values.

B. Electrical Characteristics (Dark Currents)

The current as a function of voltage characteristics for off-stoichiometric oxides incorporated into DEIS stacks with Si-rich SiO_2 injectors (see Fig. 1d) are illustrated in Fig. 7 for CVD deposited oxides with N_2O to SiH_4 ratios of $R_0=200$, 50, 40 and 30. This ratio $R_0=200$ corresponds to a stoichiometric oxide while $R_0=30$, 40 and 50 corresponds to approximately 5-6%, 3-4%, 1-2% excess silicon in the oxide, respectively, as deduced from Rutherford backscattering (RBS) measurements.[16] The ramp I-V plots shown in Fig. 7 were generated by applying a constant rate voltage ramp ($|dV_g/dt| = 0.5$ V/sec) to the samples

up to a measured current of the order of 10^{-5} to 10^{-4} A; the voltage ramp was then reversed at the same rate. A displacement current ($I_D = C \cdot dV_g/dt$, where C is the capacitance associated with the MOS structures) shows as an approximately constant current component at low applied electric fields. The hysterisis observed in the I-V characteristics is associated with permanent charge trapping in the oxide.[31] This hysterisis did not recover after periods of days with the gate electrode floating or grounded on off-stoichiometric oxides in DEIS stacks (Fig. 7) or MOS capacitors with no injectors (Fig. 8). As the excess silicon in the oxides is increased (R_0 is decreased) two important effects are observed: i) the permanent trapped charge in the oxide decreases, and ii) lower electric fields are required to induce charge transport across the devices.

Current as a function of voltage characteristics for off-stoichiometric oxides with increasing Si content incorporated into MOS capacitors without Si-rich SiO_2 injectors (see Fig. 1a) showed a behavior similar to the data of Fig. 7, but with all curves moved to larger gate voltages. This is shown in Fig. 8. Another example of this I-V data using a point-by-point I-V technique is shown in Fig. 9 for a 1200 Å thick off-stoichiometric oxide in an MOS configuration with no injectors. In this measurement the gate voltage is incremented in 2.5 V steps. The current at each point is measured after the transient capacitive current associated with the voltage increment is negligible (~ 10 sec after the voltage step). Figure 10 shows point-by-point measurements on a 300 Å thick onf-stoichiometric oxide in an MOS configuration. The behavior of this data is similar to that in Fig. 9 only the corresponding current levels are achieved at lower voltages. The current values at low fields ($\leq$ 1 MV/cm) are smaller than 10^{-12} A which is the sensitivity limit of our picoammeter. The dark current data for the large area (.005 cm^2), Al-gated, capacitors in Fig. 9 ($R_0 = 50$, 1200 Å oxide thickness) and Fig. 10 ($R_0=50$, 300 Å oxide thickness) shows a weak dependence on gate voltage polarity (V_g^+ or V_g^-). If the dark current density ($J = I/A$) as a function of the average electric field E is plotted and compared for Figs. 9 and 10, only a weak dependence on off-stoichiometric oxide thickness is observed. The thinner oxide structure (300 Å) has a larger current density (less than a factor of 10 larger) for comparable average electric fields than the thicker oxide structure (1200 Å).

Measurements to determine the time dependence of the currents were performed on off-stoichiometric oxides (300 - 600 Å in thickness with 1% to 6%, $R_0 = 30 - 50$, excess atomic Si compositions) with or without 200 Å thick $R_0 = 3$ ($\geq$ 13% excess atomic Si) injectors. In these measurements a single voltage step was applied to the gate electrode and the current was monitored on an oscilloscope. No significant changes (less than a factor of

five) in the currents were observed for time:from 500 μsec (after the capacitive transients had decayed) to 20 sec for various step voltages which produced dark current densities from 10^{-2} to 10^{-4} A/cm^2. These data imply that the reversible space charges in the off-stoichiometric oxide and injector layers are set-up in times less than 500 μsec, and that the current measured in the external circuit is near a steady state value for times greater than this.

The electron current induced through a thin off-stoichiometric oxide layer in an MOS configuration exhibits an exponential dependence on the inverse applied electric field as in the case of electron injection or conduction by tunneling across an insulating layer.[32,33] Figure 11 illustrates the behavior of the current density over the squared average electric field (J/E^2) as a function of $1/E$ for the negative gate voltage data shown in Fig. 9. If a Fowler-Nordheim injection mechanism[32] is presumed (that is, $J = aE^2 e^{-b/E}$ where a and b are constants independent of the electric field E which describes tunneling through a triangular energy barrier), an effective barrier height $\Phi_{eff} = 0.60$ eV can be calculated from the slope of the data in this semilogarithmic plot. The details of this calculation will be discussed at the end of this section.

Although **permanent** electron trapping into "deep" sites is increasingly suppressed with increasing Si content in the off-stoichiometric oxide layer, a **reversible** space charge builds up on the Si islands near the injecting interface. This is demonstrated in Fig. 12. The flatband voltage shift in this figure is a measure of the trapped charge near the injecting interface, which is formed by the 0.2 Ωcm <100> p-type Si substrate onto which the various off-stoichiometric layers have been deposited. The carriers (electrons) are injected into the oxide by an avalanche injection technique in which the hot carriers, which can surmount or tunnel through the Si-oxide interfacial energy barrier, are created in a high field depletion region of the Si by carrier multiplication induced by voltage pulsing.[8] As can be seen from this figure the $R_o = 30$ off-stoichiometric layer of the MOS device can be charged or discharged very rapidly. Injected carriers rapidly thermalize into the potential wells of the Si islands and then move from Si island to island until they reach the metal counter-electrode. When the avalanche current is shut-off, the charge remaining on the Si islands leaks off to the contacts (probably mostly to the Si substrate which is nearest). As the excess Si content of the off-stoichiometric oxide layer is decreased, its conductivity is decreased since there are fewer Si islands to which carriers can tunnel and the amount of space charge trapped on Si islands near the injecting interface increases. Compare $R_o = 30$ and $R_o = 50$ oxide data in Fig. 12. When the avalanche current is shut-off, the discharge rate is slower on the $R_o = 50$ oxide. As the Si content is further decreased ($R_o = 80$) the charging and discharging rates dramatically slow down. When the oxide becomes stoichiometric ($R_o \geq 100$) and completely devoid of

excess Si (and therefore Si islands), the trapping rate and trapped charge level return to that normally observed for the deep sites in SiO_2 related to H_2O and to OH bonded to Si.[34]

The temperature dependence of the ramp I-V curves for off-stoichiometric oxides (R_o=50) was investigated in the range of 80°K to 572°K. Figure 13 shows the ramp I-V results for a 1200 Å R_o = 50 off-stoichiometric film incorporated into a Al-oxide-Si MOS capacitor structure under negative voltage bias for different sample temperatures. A shift towards higher current values is observed with increasing temperature. However, these variations of the I-V characteristics with temperature are similar in magnitude to those observed for stoichiometric oxides where the current measured in the external circuit is limited by Fowler-Nordheim tunneling near the injecting contact-oxide interface which is strongly dependent on the electric field.[13,32,35]

Figure 14 shows data similar to Fig. 13 only for a Au-oxide-Si MOS capacitor under negative gate voltage bias. Although these data shown a slightly larger variation of the current with temperature than seen for the Al-gated structure in Fig. 13, they are still very similar implying that there is no pronounced dependence of the I-V characteristics with metal electrode for temperatures in the range from 80 to 572°K. A temperature dependence of the I-V data similar to Figs. 13 and 14 was seen for positive gate voltages on similar capacitor structures. Also for other samples with varying off-stoichiometric oxide thickness (in the range from 300 Å to 1200 Å) with or without Si-rich SiO_2 injectors for either gate voltage polarity, a similar temperature dependence to that observed in Figs. 13 and 14 was seen for current density as a function of average electric field characteristics. Figure 15 presents data like that in Fig. 13 replotted to show that the temperature dependence of the dark currents in the off-stoichiometric oxide film increases with decreasing average electric field.

The R_o = 50 oxide layers for the MOS structures used for obtaining the data in Figs. 13-15 were 1200 Å thick. As the film thickness is decreased to 300 Å a slight dependence on metal electrode (Au gives less current than Al for V_g^-) was observed on several capacitor runs. This ms consistent with the electronic space charge layer (electrons) in the vicinity of the contact-oxide interface not completely screening the effect of the metal work function (therefore interfacial energy barrier height) differences for negative gate bias. For positive gate bias V_g^+, no effect of the metal contact was seen even for R_o = 50 oxides 300 Å in thickness. This implies that holes are not the dominant carrier in these films.[27] This point will be discussed in detail in section III-F using carrier separation techniques.

The dependence of off-stoichiometric oxide current as a function of average electric field for either gate voltage polarity (V_g^+ or V_g^-) also showed no strong dependence on single crystal Si substrate resistivity in the range from 2 Ωcm to .001 Ωcm for either p-type or n-type wafers. For 1000 Å thick R_o = 50 oxides in an MOS configuration, the largest effect was seen for positive gate voltage (V_g^+) with the .001 Ωcm p-type Si substrates. A decrease in the current measured in the external circuit by less than a factor of 10 for moderate electric fields as compared to negative gate voltage (V_g^-) on this same wafer or for either voltage polarity on the other wafers (.001 Ωcm or 2 Ωcm n-type, 2 Ωcm p-type) was observed.

Figure 16 shows the temperature dependence at a constant average electric field magnitude of 2.5 MV/cm under negative voltage bias on Al-gated MOS capacitors with varying thicknesses (300 Å - 1200 Å) of off-stoichiometric (R_o = 50) SiO_2 layers. As seen from these data, there is no pronounced dependence of the temperature variation of the conductivity of the off-stoichiometric oxides with thickness. Also, this figure shows no strong dependence of the current with thickness (less than a factor of 10 change) for a given average electric field. This weak dependence of the current on off-stoichiometric oxide thickness was observed for both voltage polarities ($V_g^\pm$) over the range of average electric fields ($\sim$ .5 to 6 MV/cm) and film thickness (150 Å to 1200 Å) studied here. It was also observed for structures with or without Si-rich SiO_2 injectors for off-stoichiometric oxide compositions from R_o = 50 (1-2% excess atomic Si) to R_o = 30 (5-6% excess atomic Si).

The dependence on area and perimeter of current density J as a function of average electric field E for a wide variety of different off-stoichiometric oxide capacitors (see Fig. 1a) and FETs (see Fig. 2b) was investigated on structures from the MDT-HOTEL 1 series of wafers for 300 Å thick R_o = 50 (wafer G) and 600 Å thick R_o = 50 (wafer H) oxides. The areas A were varied from 6.5 x 10^{-4} cm^2 to 4 x 10^{-7} cm^2 and the perimeters P were varied from 1 x 10^{-1} cm to 3 x 10^{-3} cm. Correspondingly, the area to perimeter ratios A/P were varied from 6.4 x 10^{-3} cm to 1.3 x 10^{-4} cm. No dependence of J vs. E on A, P, or A/P was seen for either gate voltage polarity (V_g^+ or V_g^-) on these devices from wafers MDT-HOTEL 1-G or H. This further implies that the dark current observed for off-stoichiometric oxides are microscopically controlled and not associated with oxide defects or edges of contacting electrodes. Also, no pronounced dependence was seen for these J vs. E data on gate voltage polarity or off-stoichiometric oxide thickness which is consistent with data for the larger area capacitors (.005 cm^2) discussed previously.

The effects on the current-voltage characteristics of DEIS stacks incorporating off-stoichiometric oxides as compared to DEIS stacks with stoichiometric intervening oxides

for both polarities is shown in Fig. 17. The reduction in charge trapping seen in Fig. 7 for negative gate bias V_g^- is also observed for positive gate bias V_g^+ on off-stoichiometric oxide films as shown in this figure. Figure 18 compares DEIS stacks with off-stoichiometric oxides in an MOS capacitor configuration (see Fig. 1d) with MOS capacitors without the Si-rich SiO_2 injector layers of the DEIS (see Fig. 1a). As seen in this figure, the DEIS current-voltage characteristic is moved to lower gate voltage magnitudes as compared to the off-stoichiometric oxide alone in an MOS configuration. This implies the DEIS structure with the intervening off-stoichiometric oxide is still somewhat sensitive to the localized electric field enhancement near the Si-rich SiO_2 injector interface with the intervening oxide layer, and that these interfaces must influence and limit the currents measured in the external circuit somewhat.[13]

Figures 19 and 20 show J/E^2 as a function of $1/E$ for DEIS structures with off-stoichiometric oxides ($R_o = 50$) of various thickness (300 - 1200 Å) for negative and positive gate voltage bias, respectively. Again as seen in Fig. 16 and discussed previously for off-stoichiometric oxides alone, the current-field dependence is not a strong function of oxide thickness for these structures. The solid lines in these figures are fit using least squares to the Fowler-Nordheim relationship $J = aE^2 e^{-b/E}$ where Φ_{eff} is determined from b through the relationship

$$\Phi_{eff} = \left[\frac{3\hbar q\, b}{4(2m^*)^{1/2}} \right]^{2/3} \tag{3}$$

where $\hbar$ is Planck's constant divided by 2π (1.055×10^{-34} joule-sec) and m^* is the effective mass of a tunneling electron (≈ 0.5 of the free electron mass[32]). The values of Φ_{eff} are $\sim .4$ eV for either polarity as expected from the data of Figs. 17-18 where little, if any, polarity dependence is observed. The effective barrier $\Phi_{eff} = 0.6$ eV was deduced for I-V characteristics for a 1200 Å thick $R_o = 50$ layer in an Al-gated MOS capacitor structure (no Si-rich SiO_2 injectors) as was discussed previously (see Fig. 11). Since no strong temperature dependence was observed on structures without (as seen in Figs. 13-16) or with Si-rich SiO_2 injectors, these effective barriers $\Phi_{eff} \sim .4$ eV (with injectors) and $\Phi_{eff} \sim .6$ eV (without injectors) cannot be true energy barrier heights. Since field enhancement produced near the injector-oxide interface is still present (as shown in Fig. 18), it seems as though $\Phi_{eff} \sim .4$ eV (with injectors) is influenced somewhat by local field enhancements near this interface. Capacitor structures with just the off-stoichiometric oxide layer and no injectors also would be expected to have some field enhancement due to the size, shape, and density of the tiny islands in the oxide layer.[33] Assuming a field dependence given by a Fowler-Nordheim like tunneling

106

dependence, the true energy barrier height in the tunneling process would be given by

$$\Phi = x^{2.3} \Phi_{eff} \qquad (4)$$

where x is an average field enhancement factor.[1,31] A relative comparison between structures *without* and *with* injectors can be made using Eq. 4 with $\Phi = .6$ eV (no injectors) and $\Phi_{eff} = .4$ eV (with injectors). From this calculation for $R_o = 50$ oxides, $x = 1.8$ which is reasonably close to the values of 1.5 to 2 observed for the average electric field enhancement near Si-rich SiO_2 injector interfaces with stoichiometric SiO_2 layers.[31,36] This calculated value for the average relative field enhancement (~ 1.8) is consistent with those observed (~ 1.5 to 2) by comparing the voltages of the data of Fig. 18 at a constant current level. The trends in the values of Φ_{eff} discussed above determined from fitting I-V data to a Fowler-Nordheim relationship for off-stoichiometric intervening oxides with or without injectors was observed *over a range of oxide thickness* (300 Å to 1200 Å) *and composition* ($R_o = 50$ to $R_o = 30$) with the values deduced for Φ_{eff} decreasing with increasing Si content (decreasing R_o values).

C. Capacitance Measurements and Dielectri. Constants

Capacitance measurements were used to determine the small-signal static dielectric constant K_s of the off-stoichiometric oxides discussed here. The dielectric constant had values between those found for stoichiometric SiO_2 ($K_s = 3.9$) and Si-rich SiO_2 injectors with $\geq$ 13% excess atomic Si ($K_s \approx 7.5$). Values for $R_o = 40$ to $R_o = 50$ oxide materials measured at 1 MHz with 15 mV (rms) signals on MOS capacitors biased into accumulation were in the range of $K \approx 4$ to 5. These small signal measurements (15 mV rms) were independent of accumulation voltage bias (for dark current levels from $\leq 10^{-10}$ A/cm^2 to 10^{-4} A/cm^2) and independent of measurement frequency from 1 MHz to 10 Hz.

Measurements of the real part of the index of refraction N using ellipsometry were used to determine the high frequency dielectric constant K_i where $\epsilon_i = K_i$ x (8.86 x 10^{-14} F/cm) = N^2 x (8.86 x 10^{-14} F/cm). The absorption coefficient κ of the complex refractive index $N = N(1 - i\kappa)$ was shown to be very small ($< .01$ for off-stoichiometric oxides with $R_o \geq 30$) by a variable angle of incidence technique discussed in a previous publication.[11] For $R_o \geq 30$ (5-6% excess atomis Si), K_i was ≤ 3 which is close to the value of 2.15 measured for stoichiometric SiO_2.

D. Electroluminescence

Another implication of the proposed conduction mechanism in off-stoichiometric SiO_2 is that the energy distribution of the electrons arriving at the positively biased electrode should have an average excess energy (E_{AV}), with respect to the collecting electrode Fermi level, lower than in the case of stoichiometric SiO_2 where E_{AV} is approximately equal to the difference in energy of the bottom of the SiO_2 conduction band and the collecting electrode Fermi level. The excess energy of these electrons arriving at the counter electrode can be dissipated through a radiative process which in turn could be used to estimate the electron energy. Measurements of the light generated are facilitated if relatively large currents can be passed through the oxides. In order to get large injected current into oxide layers at moderate to high average electric fields without destructive breakdown,[15] SEIS structures as depicted in Fig. 1c and 21 have been used. Low field breakdowns which are usually related to the contact-oxide interface are eliminated by the field screening action of the Si-rich SiO_2 injector.[15] With the SEIS under positive gate voltage bias, electrons are injected into the stoichiometric SiO_2 conduction band at thermal energies by field-enhanced Fowler-Nordheim tunneling from the Si-islands in the Si-rich SiO_2 injector layer.[13-15,31] For the MSEIS, electrons are believed to be injected directly from the Si-rich SiO_2 injector layer ($\geq$ 13% excess atomic Si) onto Si islands in the off-stoichiometric oxide layer ($\leq$ 6% excess atomic Si) from which the electrons can move across the oxide mostly by direct tunneling from Si island to Si island. These two cases (SEIS and MSEIS) are illustrated in Fig. 21 for a structure with a thin Au metal gate electrode. Luminescence emission for SEIS and MSEIS structures under positive gate bias as a function of photon energy are shown in Fig. 22 for thin circular Au gates (250 Å in thickness and .01 cm^2 in area) at an average current of 5 x 10^{-6} A. Results similar to Fig. 22 were seen up to average current levels of 1 x 10^{-4} A with the amplitude of the spectra scaling with current level after electric field dependent effects were taken into account. For stoichiometric SiO_2, a ... -off in the luminescence spectrum at $\approx$ 4.1 eV would be expected (if there is no electron heating as the carriers traverse the oxide layer) which is consistent with the interfacial energy barrier height (from the top of the metal Fermi level to the bottom of the SiO_2 conduction band) measured using internal photoemission from Au into SiO_2.[2,3] Most of the luminescence spectrum shown in Fig. 22 for the SiO_2 structure is believed to be associated with Au surface plasmons or interband transitions created by the incoming SiO_2 conduction band electrons, most of which enter the thin metal ($\leq$ 250 Å) layer with $\approx$ 4.1 eV excess energy.[37] Surface plasmons in the metal electrode, stimulated by the incoming electrons with excess energy due to the barrier step at the metal-oxide interface, can be effectively coupled out as light only in the presence of surface roughness. This roughness in

SEIS and MSEIS structures comes from the bottom Si-rich SiO_2 injector layer which will be discussed in Section III-E, and it can be further enhanced by deposition of the SEIS stacks on purposely roughened substrates.[37] For constant current conditions this Au luminescence spectrum was relatively insensitive to changes in the internal electric fields of the SiO_2 layer caused by electronic charge trapping into the "deep" H_2O or OH related sites particularly for fields ≤ 8 MV/cm. This Au luminescence spectrum would not be expected for comparable current injection under negative gate voltage, but these experiments were difficult to perform because of voltage breakdown at comparable current levels when no Si-rich SiO_2 injector is present to field-screen Au-SiO_2 contact irregularities. In Fig. 22 luminescence spectrum for the MSEIS which incorporates an off-stoichiometric oxide layer ($\leq 6\%$ excess atomic Si) shows very little light output in the blue region of the spectrum from the Au electrode. This result is consistent with the proposed mechanism for conduction in the off-stoichiometric oxide films (tunneling from Si island to Si island) where most of the incoming electrons at the metal-oxide interface would have only ≈ 1.1 eV excess potential energy (which is the energy difference of the Au-SiO_2 barrier height, 4.1 eV, and the depth of the potential well of a Si island, ≈ 3 eV).

The small amount of light output (mostly at red wavelengths, see Fig. 22) from the MSEIS was also observed to fairly independent of bias voltage polarity (V_g^+ or V_g^-) and the presence or absence of the single bottom Si-rich SiO_2 ($R_0 = 3$) injector layer. This implies that this luminescence is from the off-stoichiometric oxide itself since for V_g^+ electrons enter the metal electrode while for V_g^- they leave, and since without a bottom injector layer the metal surface is relatively smooth and lacks the necessary roughness to effectively couple light out from surface plasmons (see Section III-E). Some detailed spectral differences were seen on MSEISs with different metal gate electrodes (Al or Au) due to differences in the metal transmission properties and the possibility of more light output with Au electrodes below $\approx$ 1.1 eV due to surface plasmons or interband transitions, as discussed previously. The detailed structure (peaks and valleys in the spectrum) of this light output was very different from that due to the light emission from the Au or Al layers observed only under positive voltage polarity (V_g^+) on SEISs with stoichiomteric oxide layers. A strong increase in the amplitude of this luminescence spectrum was observed when the off-stoichiometric oxide thickness was increased from 500 Å to 1000 Å with the average field and average current held at similar values. This increase in light output is not due to optical interference effects, since the peak and valley positions in the luminescence spectrum did not change as the oxide thickness was increased from 500 Å to 1000 Å. Smaller differences in the luminescence amplitude on stoichiometric oxide SEISs with 500 Å and 1000 Å oxide layers where surface plasmon

luminescence and metal interband transitions dominate were seen under similar conditions. Also, the luminescence from the off-stoichiometric oxides exhibited an electric field dependence, with the various spectral peaks being effected differently by the field. Most of the red portion of this luminescence spectrum (≤ 2 eV) is believed to be due to bandgap or near bandgap recombination processes on the Si islands,[38] particularly near the positively biased contact where a reversible positive space charge has been observed to build-up. This positive space charge could be in part due to holes even through holes were not observed to be the dominant carrier for these materials as will be discussed in section III-F. A portion of this red spectrum, particularly for Au (which has an ~ 1 eV larger work function than Al and therefore an ~ 1 eV larger interfacial energy barrier with SiO_2) at energies < 2 eV, could be due to some electrons entering the Au from the conduction band of a Si island with enough excess energy (see Fig. 21b) to excite some metal surface plasmons or interband transitions.

Thickness undulations (roughness) which will be discussed in section III-E could be argued to account for the current-voltage characteristics of off-stoichiometric oxides (section III-B) by a mechanism in which electrons are injected locally at the thin regions via Fowler-Nordheim tunneling into the SiO_2 conduction band. However, for the same average current per unit area, a luminescence spectrum comparable to that for the stoichiometric oxide film in Fig. 22 would be expected. Also current density-electric field characteristics would be expected to be a strong function of off-stoichiometric oxide thickness which was not observed in the thickness range from 300 Å to 1200 Å (see section III-B).

E. Surface Roughness

Scanning electron micrographs (SEMs) were taken on the top surfaces of some of the various oxide layers used in the capacitors and devices discussed here. SEMs were performed using an electron beam energy of 35 KV with the samples tilted at a $45°$ angle with respect to the incident electron beam direction. All oxide layers were deposited on, 2 Ωcm. <100> orientation, n-type, single crystal Si substrates which were planar with respect to the resolution of the SEM measurements. All films had a 1000°C anneal in N_2 for 30 minutes after deposition to remove H_2O and densify the films, as was similarly done on all devices and capacitors used in this study. To minimize charge build-up on the oxide surfaces during the SEM measurements, all samples were coated with 300 Å thick Au films. Figure 23 compares 500 Å thick CVD stoichiometric ($R_0 = 200$) and off-stoichiometric ($R_0 = 30$ which has $\approx$ 5-6% excess atomic Si) oxide films. The off-stoichiometric oxide in Fig. 23b with $R_0 = 30$ has some slight amount of roughness on its top surface (the largest hillocks are ≤ 200 Å high

with $\leq$ 400 Å bases). Some nearly smooth surfaces were observed on other off-stoichiometric oxide films with $R_o = 50$ ($\approx$ 1-2% excess atomic Si), even up to 1200 Å in thickness. The stoichiometric oxide surface in Fig. 23a is smooth. It is interesting to note that off-stoichiometric oxide layers of thickness $\leq$ 150 Å were usually leaky or broke-down destructively easily at low bias voltages. Figure 24 compares SEMs of stacked layers of CVD Si-rich SiO_2 (with $\geq$ 13% excess atomic) deposited between the Si substrate and the thicker 500 Å CVD layer of either stoichiometric or off-stoichiometric oxide ($R_o = 30$). This stack is comparable to a single electron injector structure (SEIS). The density of hillocks in Fig. 24 is much greater than in Fig. 23b and is consistent with what has been reported previously for deposited Si-rich SiO_2 films on single crystal Si substrates.[15] This roughness increases in size as the Si-rich SiO_2 layers are made thicker, but the shape remains $\approx$ hemispherical.[15] Although this roughness could contribute to the injection of carriers from either a top metal contact interface with an off-stoichiometric oxide film or from an injector-oxide or injector—off-stoichiometric oxide interface, there is no correlation of this large scale roughness with the magnitude of the current injection[15,39] (probably due to conformal mapping of the thin oxide layers). This was seen previously from the invariance of ramped I-V data for DEIS stacks with stoichiometric intervening oxide layers (400 Å thick) but with Si-rich SiO_2 injectors of varying thickness from 100 Å to 1000 Å.[39] Results presented in section III-B (See Figs. 16, 19, 20) for off-stoichiometric films of various thickness (150 Å - 1200 Å) and composition ($R_o = 50$ to 30) in MOS or DEIS stack capacitors also show little effect of the large scale roughness on current density - average electric field characteristics for either voltage polarity. However, electric field distortion from roughness on a much finer scale ($\leq$ 50 Å) is believed to influence the conduction characteristics of off-stoichiometric layers when they are used in DEIS stacks as the intervening oxide. This roughness is due to the injecting Si islands ($\leq$ 50 Å in size) in the Si-rich SiO_2 ($\geq$ 13% excess atomic Si) injectors of the DEIS and will be discussed in section IV.

F. Hole Conduction

The results in section III-A through III-D imply that electrons and not holes are the dominant carrier in the off-stoichiometric oxide films. To obtain a more direct measure of the relative ratio of electron currents to hole currents, the transistor structures shown in Fig. 2b under positive gate voltage bias were used to separate the motion of the carriers. In these structures, holes flow across the oxide to the p-type Si substrate after injection from the poly-Si gate electrode, while electrons flow across the oxide to the poly-Si gate after injection from the surface inversion layer of electrons supplied by the source and drain diffusions. All

measurements were performed in the dark using the circuit shown in Fig. 2b. The gate voltage was ramped at a constant rate of 0.5 V/sec with all currents being simultaneously recorded by identical log-picoameters. This technique for separating charge carrier flow in an insulating layer, originally developed by Weinberg et al.[40] and Ginovker et al.,[41] independently, has been used previously with stoichiometric SiO_2 and Si_3N_4 layers by several investigators.[40-43] Electrons were found to be the dominant carrier in SiO_2 with a small value for the ratio $\alpha = I_p/I_n$ of the substrate to diffusion current (see Fig. 2b) for positive gate voltage polarity.[40] However in Si_3N_4, holes were found by most investigators to be the dominant carrier with a ratio of $\alpha \approx .5$ for positive gate voltage.[42,43] Data in Fig. 25 shows α for various SiO_2 and off-stoichiometric oxide ($R_o = 50$) combinations used as the gate insulator in the n-channel FETs depicted in Fig. 2b. As seen in this figure, $\alpha \leq 3 \times 10^{-3}$ or $\alpha \leq 1 \times 10^{-2}$ for the off-stoichiometric oxide or stoichiometric thermal SiO_2 structures, respectively, under positive gate voltage bias. The sum of the substrate and diffusion currents always equaled the gate current ($I_n + I_p = I$, see Fig. 2b) for all devices in Fig. 25. The ratio of hole to electron currents in Fig. 25 for devices with off-stoichiometric ($R_o = 50$) oxide layers is approximately constant for a wide range of average electric field variation across the gate insulator. This observation seems hard to account for if I_p is really a measure of the hole current injected from the gate electrode and conducted through the gate insulator because different energy barriers for hole injection and conduction are expected compared to that for electrons.[42] This will be discussed in more detail in section IV. The possibility that holes strongly recombine with electrons on Si-islands in the bulk of the oxide or at the substrate Si-oxide interface is a possibility in explaining the approximately constant value of α over such a large electric field range.[42] Few holes would be expected to recombine with electrons in the thin (≤ 100 Å) inversion layer in the Si near the interface.[40] However, it seems unlikely that bulk oxide or interface recombination would result in the small values of α observed ($\alpha \leq 3 \times 10^{-3}$). It is therefore concluded that I_p is probably a measure of a substrate current component of the injected electron current and that any component of I_p due to hole conduction is at least a factor $\leq 3 \times 10^{-3}$ of that observed for electronic motion.

Capacitance as a function of gate voltage measurements on capacitor structures similar to the devices used for the data in Fig. 25 were also performed. For structures with just an off-stoichiometric oxide layer between the Si substrate and the gate electrode, positive gate voltage or negative gate voltage stressing produced shifts of the capacitance-voltage (C-V) characteristic to larger or smaller gate voltages consistent with negative or positive charges near the substrate Si-oxide interface, respectively. However for moderate negative gate voltage stressing, the addition of a thin (70 Å thick) blocking stoichiometric SiO_2 layer in

between the Si substrate and the thicker off-stoichiometric oxide layer (See Fig. 2b) caused the C-V characteristic to shift in the direction of stored negative charges (electrons). These experiments imply that the positively stored charges seen for negative bias stressing on metal/off-stoichiometric oxide/silicon MOS capacitors originate from charge exchange between the Si substrate energy bands and the off-stoichiometric oxide potential energy wells caused by the Si islands very near ($\lesssim$ 50 Å) the substrate Si-oxide interface, in much the same way as interface states are observed with stoichiometric SiO_2 MOS structures.

IV Discussion of Results

The data presented in section III from a variety of experiments (photoconduction, dark conduction, flatband voltage, Raman spectroscopy, carrier separation using n-channel FETS, and luminescence) lead to the following conclusions about the conduction mechanism in the off-stoichiometric oxides:

a) very tiny ($\lesssim$ 30 Å) Si islands are present in the oxide films for $R_0 < 100$ ($\approx$ 1-6% excess atomic Si);

b) These islands modify the conduction mechanism; in particular, electrons appear to be predominately moving from the metal contact, Si substrate, or Si-rich SiO_2 injector interface with the oxide, into potential wells created by the Si islands rather than into the oxide conduction band. This is depicted schematically in Fig. 26;

c) The strong electric field dependence observed in the external circuit is consistent with a tunneling mechanism. If a Fowler-Nordheim-like tunneling mechanism (that is, tunneling through a triangular energy barrier) is assumed, the effective energy barriers deduced would indicate a stronger temperature dependence than observed for the dark current and photocurrents;

d) The weak dependence of the photocurrents or dark currents, measured in the external circuit, on the injecting contact-oxide interface (oxide-Si, Al, Au) suggests that a space charge layer builds up in the oxide near the injecting interface and limits the current injection into the bulk of the off-stoichiometric oxide. This space charge layer is mostly reversible and is thought to build up on the small Si islands;

e) The lack of a strong dependence of the current density as a function of average electric field, on oxide thickness in the range from 150 - 1200 Å suggests that this limiting space charge layer is less than $\sim$ 150 Å in extent.[27] This is further supported

113

by the experimental observation that some localized electric field enhancement is still observed at the injector-oxide interfaces when the off-stoichiometric oxides are incorporated into DEIS stacks with Si-rich SiO_2 injectors (see Fig. 18). This implies that the space charge layer cannot totally screen the enhanced local fields created by the curvature of the somewhat larger Si islands in the injector layer;

f)   Hole injection into and conduction through the off-stoichiometric oxides could also influence the observed currents somewhat. However, from the experimental data presented here there is no strong support for hole injection and conduction as the dominant mechanism.

Abeles and coworkers have treated the problem of carrier transport in cermets which are metal particles such as Ni, Pt or Au embedded in insulators such as SiO_2 or Al_2O_3.[33] The conductivity is assumed to be controlled by direct tunneling from metal grain to metal grain with some consideration for electric field distortion caused by the size, shape, density, and spacing of the metal grains in the oxide matrix.[33] They show that the conductivity has a low electric field region which is sensitive to temperature but not field, and a high electric field region which is sensitive to field but not temperature. The high field region is defined by

$$E \,>\,> kT/qw \tag{5}$$

where $w = s + d$ (s = metal particle separation and d = metal particle size) is a measure of the distance between equipotential surfaces in a granular metal (that is, the average distance between planes of neighboring metal grains). In the high field region the current density was shown by Abeles $et\ al.$ to be

$$J \propto E\, e^{-\frac{E_o}{E}} \tag{6}$$

where E_o is a constant. The functional form of Eq. 6 is almost identical to Fowler-Nordheim tunneling from near the metal Fermi level through a triangular barrier into the bottom of the oxide conduction band (i.e., $J = a\,E^2\,e^{-b/E}$ as discussed in section III-B).

The conductivity in off-stoichiometric oxides (basically, a Si and SiO_2 cermet) can be interpreted in terms of the high field region model of Abeles $et\ al.$ If reasonable values for w are chosen in the range from 10 Å to 50 Å, kT/qw at room temperature is between 2.5×10^5 and 5×10^4 V/cm which is much less than any of the average fields for the data presented in this study (for example, see Figs. 9 and 10). Fowler-Nordheim tunneling into the SiO_2 conduction band, as discussed previously, seems a doubtful mechanism for these off-

114

stoichiometric oxides because of the low energy barriers ($\leq$ .6 eV) deduced from data fits (see section III-B), the lack of permanent charge trapping (see section III-B), and the lack of luminescence from the metal gate electrode caused by entering SiO_2 conduction band electrons (see Section III-D). For Fowler-Nordheim tunneling to dominate at the high fields, $\Phi/q \approx 3$ V must be dropped across a distance less than the average Si island spacing s, where it has been assumed that a 3 eV energy barrier exists between the bottom of the conduction band of a Si island and the bottom of the SiO_2 conduction band (see section III-A). For a reasonable range of values for s from 10 Å to 50 Å, $E = \Phi/qs$ is between 3×10^7 V/cm and 6×10^6 V/cm which is larger than any of the average electric fields for the data presented here. Electron Fowler-Nordheim tunneling from the top of the valence band of a Si island to the bottom of the SiO_2 conduction band is even less probable since $\Phi/q \approx 4$ V and would require electric fields of 4×10^7 V/cm and 8×10^6 V/cm.

In most of the discussion, it has been assumed that the measured currents are dominated by electrons on Si islands tunneling directly from near the bottom of the Si island conduction band into an empty conduction band state on neighboring Si islands (see Fig. 26). At larger electric fields, electrons could tunnel from near the top of the Si island valence band ($\sim$ 4 eV deep from the bottom of the SiO_2 conduction band) on one Si island into empty conduction band states on another island ($\sim$ 3 eV deep from the bottom of the SiO_2 conduction band). However, this process would require fields as large as 1×10^7 V/cm if $\sim$ 1 eV is dropped between the Si islands which are assumed to be $\approx$ 10 Å apart, and it would have a reduced tunneling probability because of the larger energy barrier (4 eV compared to 3 eV) involved.

In the low field region, Abeles's model predicts

$$J \propto e^{-2\sqrt{\frac{c}{kT}}} \tag{7}$$

where c is a constant. For temperatures varying between 100°K and 400°K, Abeles *et al.* show for Ni-SiO_2 and Pt-SiO_2 cermet films with low ratios of the volume fraction of metal ($\leq$ .05) a resistivity change by a factor $\approx 10^5$ in the low field region.[33] They also show that the strong temperature dependence diminishes on Ni-SiO_2 cermets with a .08 volume fraction of metal at an average electric field of $\approx 5 \times 10^5$ V/cm where only a two order of magnitude change in the resistivity is observed in going from 1.4°K to 300°K. The data presented in Figs. 13-16 where only one to three orders of magnitude change in J is observed over a temperature range from 80°K to 572°K for average electric fields $> 1 \times 10^6$ V/cm is consistent with the model used for cermets by Abeles *et al.* . Also as seen in Figs. 13-15 the

temperature dependence of the current decreases with increasing field. As mentioned earlier in section III-B and seen in Fig. 15, the temperature dependence for the off-stoichiometric oxide conductivity is similar to that observed for stoichiometric oxides in contact with planar metal or semiconductor (Si) layers[32,35] or Si-rich SiO_2 injectors ($\geq$ 13% excess atomis Si).[13] The temperature dependence for the contact-limited electron injection observed with metal or Si - SiO_2 systems increases at lower average electric fields in a similar manner to that seen for the data in Fig. 15 for R_o = 50 off-stoichiometric oxide MOS structures.

The dark currents appear to have as exponential temperature dependence (that is, $\ell n\, J \propto T$) as seen in Figs. 15 and 16. The data in these figures do not follow a $\ell n\, J \propto -1/\sqrt{T}$ dependence (as described by Eq. 7 for cermets) very well over the entire temperature range studied ($77°K$ to $300°C$). Also, these data did not give a good fit to a single straight line when least square fit to a Poole-Frenkel type relationship (that is, $\ell n\, J \propto -(T)^{-1}$). Any type of thermally activated conduction mechanism such as Poole-Frenkel (activation from traps) where $J \alpha\, e^{-\Phi/kT}$ or Schottky emission (activation from metallic or semiconducting contacts) where $J \alpha\, T^2 e^{-\Phi/kT}$ would have a much stronger temperature dependence than observed here (see Figs. 13-15) in going from $80°K$ to $572°K$. For example, even if $\Phi \geq .1$ eV, J would increase by a factor of $\geq 10^5$ for Poole-Frenkel and $\geq 10^7$ for Schottky emission over this temperature range. For larger energy barriers, the increase would even be more. Poole-Frenkel conduction is a field-assisted, thermally-activated process where the currents measured in the external circuit are limited by carrier excitation from trapping levels to a band edge in the bulk of the insulator.[27] Here again, injection from the contacts is limited under steady state conditions by the trapped space charge near the injecting interfaces and the bulk conductivity of the material. Poole-Frenkel conduction is thought to describe the current-voltage characteristics of thick ($\geq$ 1000 Å) CVD Si_3N_4 (an $\approx$ 5.2 eV band-gap insulator) films where both the motion of electrons and holes and their subsequent trapping and detrapping are important.[27]

The dominant term controlling the field dependence of Poole-Frenkel conduction comes from energy barrier lowering where $\Phi = \Phi_0 - \Delta\Phi_{PF}$ with $\Delta\Phi_{PF} = q\,(q\,E/\pi\epsilon_i)^{1/2}$ as discussed in section III-A. Therefore, it can be easily shown that $J \alpha\, E\, e^{\beta E^{1/2}}$ where $\beta = q\,(q/\pi\epsilon_i)^{1/2}/kT$.[30] Least squares fits of J vs E data (such as that in Figs. 9,10,13,14,17 and 18) for off-stoichiometric films (R_o = 50 to 30) to this Poole Frenkel relationship were performed to deduce values of ϵ_i. These numbers deduced for ϵ_i gave values for the high frequency dielectric constant K_i in the range from 5 to 13 which are unrealistically larger than the actual measured values ≤ 3 (see section III-C). If localized, field enhancement effects were taken into account (that is the local electric fields near the Si islands controlling the

measured current are larger than the average electric field E), the values of K_i deduced from this data fit would be even larger. Poole-Frenkel conduction also seems a doubtful mechanism to describe the currents measured for the off-stoichiometric oxide films because of the large energy barriers (3 to 4 eV) measured for the optical activation of electrons trapped on Si islands to the SiO_2 conduction band discussed in section III-A and because of the lack of luminescence from the metal gate electrode caused by entering SiO_2 conduction band electrons discussed in section III-D.

Recently Ni and Arnold[44] have shown experimentally a low-field temperature dependence that follows a relationship given by $J \sim T e^{-(E_{g_{Si}}/2kT)}$ where $E_{g_{Si}}$ is the bandgap energy of the Si particles in Si-rich SiO_2 layers with >50% atomic Si concentrations. This model assumes that at low electric fields conduction in these films is dominated by a bulk limited process in which carriers (mostly electrons) thermally generated within the Si grains tunnel through energy barriers formed by intergrain SiO_2 layers. $E_{g_{Si}} \sim 1.1$ eV is deduced from the conductivity vs. $1/T$ data for their films containing 81% atomic Si.[44] The films studied here have much less atomic Si ($\leq$ 39% atomic Si) and were studied at larger average electric fields (> 1 MV/cm). J does not follow a $T e^{-(E_{g_{Si}}/2kT)}$ dependence over the entire range of temperatures (77°K to 300°C). Fitting any portion of the current data such as that in Figs. 15 and 16 to this relationship of Ni and Arnold, yielded low activation energies (the largest value deduced was $E_{g_{Si}} \sim .7$ eV from the -12 V data of Fig. 15) consistent with the weak temperature dependence observed experimentally for electric fields > 1 MV/cm. Also, currents measured in the external circuit at these electric fields (> 1 MV/cm) could not be enhanced by intense monochromatic light from a He-Ne laser (1.96 eV energy) which was used to increase the number of carriers supplied by bandgap ionization processes of the Si islands in the off-stoichiometric oxides ($\leq$ 39% atomic Si).

The discussion has ignored the contribution of hole injection and conduction in the off-stoichiometric oxide materials. The experimental results presented in section III-F together with those from section III-A and B strongly imply that electron conduction dominates the observed current. This is also consistent with simple energy barrier considerations. Hole injection from near the top of the Fermi level for Al and Au or from near the top of the valence band edge for Si (either the substrate or a Si island) into the valence band associated with a Si island would "see" an $\approx$ 6 eV (Al), 5 eV (Au), or 5 eV (Si) energy barrier assuming the band-gap of SiO_2 is $\approx$ 9 eV. Electron injection from near the top of the Fermi level for Al or Au or near the bottom of the conduction band edge for Si (either the substrate or a Si island) into the conduction band associated with a Si island would be limited by smaller energy barriers of $\approx$ 3 eV (Al), 4 eV (Au), and 3 eV (Si). Since conduction mechanisms involving

either direct or Fowler-Nordheim tunneling are strongly dependent on energy barrier heights, this implies that electron conduction should dominate in these off-stoichiometric oxides, as has been observed.

For 1% to 6% excess atomic Si used for the off-stoichiometric oxide layers, the average Si islands densities are in the range from $\sim 10^{20}$ to 10^{21} cm^{-3} for 5 Å diameter islands ($\sim$3 atoms of Si), in the range from $\sim 10^{19}$ to 10^{20} cm^{-3} for 10 Å diameter islands ($\sim$26 atoms of Si), or in the range from $\sim 10^{18}$ to 10^{19} cm^{-3} for 20 Å diameter islands ($\sim$209 atoms of Si). These densities n_0 would correspond to an average Si island separation s of 17 to 5 Å for densities of 10^{20} to 10^{21} cm^{-3} with d = 5 Å, 36 to 12 Å for densities of 10^{19} to 10^{20} cm^{-3} with d = 10 Å, and 80 to 26 Å for densities of 10^{18} to 10^{19} cm^{-3} with d = 20 Å where it has been assumed that s = $(n_0)^{-1/3}$ - d. In all cases except for d = 20 Å with $n_0 = 10^{18}$ cm^{-3}, the Si islands are not separated by more than 36 Å which is consistent with tunneling distances expected for measurable current flow through 3-4 eV energy barriers over the electric field range of interest. For Si island separations $\leq$ 50 Å, coulombic fields from two adjacent sites could overlap. This overlap would reduce a normally 3 eV energy barrier to $\leq$ .3 eV for s $\leq$ 10 Å over the electric field range of interest.[30] However, for this to occur would require that the Si islands maintain a net positive space charge everywhere in the film if electrons are the dominant charge carriers.[30] Clearly from the photoconductivity and photodetrapping results in section III-A, the dominant energy barriers are $\geq$ 3 eV. Also from the results in section III, the off-stoichiometric oxide films have a mostly reversible negative space charge near the cathode which controls the dark currents measured in the external circuit. Although some positive space charge can be observed near the anode (possibly due to ionized donor-like impurities or holes in the Si island regions), it is doubtful from the experimental evidence presented in Section III that this can significantly control the external circuit currents.

Si islands $\leq$ 5 Å (< 3 atoms of Si) in diameter seem unlikely as the dominant particle size. The photoconductivity and photodetrapping experiments discussed in section III-A (see Figs. 4-6) showed two optical thresholds at $\sim$ 3 eV and 4 eV characteristic of bulk Si (more than 1 or 2 Si atoms) in contact with SiO_2. The Raman spectra (Fig. 2) and XPS measurements discussed in section III-A are also consistent with most of the Si islands being large enough (> 5 Å) to behave like bulk Si. The limiting case of a cluster of 2 Si atoms should correspond to an oxygen vacancy in SiO_2. Calculations for the ideal (no-relaxation) oxygen vacancy in α-quartz gives two levels in the SiO_2 energy gap separated by $\sim$ 4 eV.[45] These two levels correspond to the bonding and antibonding state formed by the two Si dangling bonds. Also for Si island sizes $\leq$ 10 Å, "particle-in-a-box" type quantum effects could force the lowest possible total energy of a charge carrier in a Si region to be $\geq$ .4 eV with respect to a

Si band edge assuming the mass of the carrier is equivalent to that of an electron in free space.[46,47] For example, an electron would always have $\sim$ .4 eV excess energy, even at T = $0°$K, with respect to the bottom of the conduction band of a Si region if that region were $\sim$ 10 Å in size. These effects, if present, might be detected by experiments such as electroluminescence (see section III-D) or photoluminescence[46] which probe the physical nature of the Si islands themselves.

Most of the previous discussion concerning electric fields in the off-stoichiometric oxides has been in terms of the average electric field [i.e., $E = (V_g - \phi_{ms} - \Psi_s)/ \ell_o$]. The details of the microscopic fields between the tiny Si islands and how they are influenced by the size, shape, and density of the islands is clearly beyond the scope of this study. Detailed treatments of the randomness in size, shape, and separation of the Si particle distribution and the interactions of the electric fields between neighboring particles is beyond the scope of this study. Also, if the electric fields are really successfully screened by the Si islands (a net negligible voltage drop across the islands) a larger fraction of the applied voltage should be dropped across the stoichiometric oxide poprtions of these Si/SiO$_2$ cermets.

A simple model based on quantum mechanical tunneling of a thermal carrier through a rectangular potential barrier which has a voltage Δv dropped across it can be shown to yield the significant exponential terms in the Abeles high field result (see Eq. 6) and connect the actual energy barrier heights Φ ($\sim$ 3 eV for Si island conduction band bottom to SiO$_2$ conduction band bottom, see Figs. 21 and 26) with the effective values Φ_{eff} (such as 0.6 eV) deduced from the data fitting procedures on current-voltage characteristics described in section III-B. A WKB approximation (that is, a high, broad, and smoothly varying potential barrier) can be used to determine the transmission coefficient T (tunneling porbability) for an electron moving in the direction of higher potential (back tunneling is neglected) with near thermal energy $E_c <<$ Φ though a rectangular potential barrier with a voltage Δv dropped it[47,48] (see Fig. 26). The variation of the potential with distance Δx across the barrier would be given by

$$V(x) = \frac{\Phi}{q} - \frac{\Delta x \cdot \Delta v}{s} , \qquad (8)$$

and the corresponding relationship for the tunneling probability using the WKB approximation[47,48] would be given by

$$T = e^{-\frac{4(2m^*)^{\frac{1}{2}}}{3\hbar q \Delta v / s} \left[\Phi^{3/2} - (\Phi - q\Delta v)^{3/2}\right]} \qquad (9)$$

If the Si islands screen any fraction of the applied field or enhance the field locally, $\Delta v \, s > E$ where $E = (V_g - \phi_{ms} - \Psi_s) / \ell_o$ is the average electric field as discussed previously. Using Eq. 9, letting $\Delta v/s = \eta E$ where $\eta \geq 1$, and assuming $J \sim T$, it is easily shown

$$J \sim e^{-\frac{4(2m^*)^{1/2}}{3\hbar q E} \phi_{eff}^{3/2}}$$ (10)

where

$$\Phi_{eff} = \left[\frac{\Phi^{3/2} - (\Phi - q\Delta v)^{3/2}}{\eta} \right]^{2/3}$$ (11)

For $q \, \Delta v = \Phi$, Eq. 10 reduces to the form of the Fowler-Nordheim relationship for tunneling through a traingular energy barrier (except for the E^2 pre-exponential factor), as discussed previously (see section III-B), which in this case would be for a Si island conduction band electron tunneling into the SiO_2 conduction band (see Fig. 26). For moderate values of Δv in the range $\Phi > q \, \Delta v >> kT$, Φ_{eff} would be a weak function of E and Eq. 10 would reduce to a form similar to the Abeles high field result (except for the E pre-exponential factor). If the Si island portions of the cermets have no voltage dropped across them, it can easily be shown that $\eta = 1 + \frac{d}{s} = w/d$ assuming all islands are similar in size and equally spaced on the average,[33] and that the local bulk oxide electric field enhancement due to Si island curvature or reversible space charge build-up near the contacts is negligible. If field enhancement is not negligible, $\eta > w/d$. As the Si content of the off-stoichiometric films increases (R_0 decreases in value), $\eta = w/d$ will increase[33] and Φ_{eff} will decrease. This is consistent with the observations discussed in section III-B. For reasonable values of η (for example, 2 or 4) with $\Phi = 3$ eV, $\Phi_{eff} \leq 1$ eV for any value of $\Delta v \leq 1$ V ($\eta = 2$) or any value of $\Delta v \leq 2$ V ($\eta = 4$). This is also consistent with the results for Φ_{eff} discussed in section III-B.

By comparing the relationship for current in Eq. 10 due to tunneling with the relationship for a thermally activated mechanism such as Poole-Frenkel conduction ($J \, \alpha \, e^{-\phi/kT}$), it can be easily shown that for tunneling to dominate over thermally activated conduction

$$\frac{\Phi}{kT} > \frac{4 (2m^*)^{1/2} \Phi_{eff}^{3/2}}{3\hbar q E}$$ (12)

This inequality holds for all cases discussed here; for example, even for a worst case situation where $E = 1 \times 10^6$ V/cm, $\Phi_{eff} = 0.6$ eV, $T = 300°C$, and $\Phi = \Phi_0 - \Delta\Phi_{PF} \geq 1.2$ eV.

Another important observation, which is believed to be linked to the Si islands in the off-stoichiometric oxides and their modification of the conductivity, is the apparent decrease in

120

the permanent trapping of injected electrons into the deep (4-5 eV from the bottom of the conduction band[7]) trapping sites normally found in all stoichiometric SiO_2 layers. These sites include traps with densities $\geq 10^{11}$ cm^{-2} and capture cross sections of $\leq 10^{-17}$ cm^2. Two of these larger cross section traps ($\approx 10^{-17}$ and 10^{-18} cm^2) have been shown to be related to OH bonded to Si and H_2O.[34] These traps can be reduced by extended high temperature annealing usually at 1000°C in N_2 for times ≥ 30 min. However, there are also smaller cross-section traps $\leq 10^{-19}$ cm^2 which appear to be unrelated to OH or H_2O and which cannot be minimized by high temperature annealing or other processing variations.[8] Several possible explanations for the **minimization** of permanent trapped electron build-up on these deep sites depicted in Fig. 26, consistent with the proposed picture for the conduction mechanism and the Si inclusions in these films, are the following:

a)   Capture of an electron tunneling from Si to Si island in the off-stoichiometric oxides is expected to be different than capture of a "free" electronic carrier from the conduction band of an stoichiometric SiO_2 layer. For instance, energetically shallow levels which may participate during the initial stages of electron capture from the conduction band will not be energetically accessible to electrons tunneling through ≥ 3 eV deep wells created by the Si islands, particularly if these tunneling electrons are thermalized while moving through each Si island (that is, they are not getting "hot" with respect to a band edge). However, a certain small fraction of the injected electrons still has a finite probability of getting into the SiO_2 conduction band. A fraction of these electrons which can get into the oxide conduction band also would have a finite probability of traversing the entire oxide film without being captured into the potential well of a Si island from which the carrier could then move predominately from Si island to Si island with a small probability of being re-emitted in the SiO_2 conduction band. This small fraction of carriers that can get into the oxide conduction band and move some finite distance in the SiO_2 still have a probability of being trapped into deep states. Note some hysterisis in the ramp I-V characteristics (see Figs. 7-8) due to permanent trapping was observed on the off-stoichiometric oxides with small amounts of excess Si ($\leq 3\%$). This trend in electron trapping is also observed when the off-stoichiometric oxides in DEIS stacks are used in electrically alterable memories which will be discussed and demonstrated in the next section.

b)   Another possible explanation for the minimization of permanent trapped electronic space charge in the oxide films is the tunneling of carriers trapped on "deep" sites under energetically favorable conditions to near-by Si islands. This is depicted schematically in Fig. 26. Energetically favorable conductions would depend on the

121

applied electric field and local electric fields which would be influenced by reversible space charge distributions on Si islands, particularly near the injecting interfaces. The electric field is necessary to allow carriers which are trapped $\geq$ 4 eV from the bottom of the SiO_2 conduction band to tunnel directly into empty states associated with the Si island regions which are $\leq$ 3 eV from the bottom of the SiO_2 conduction band, assuming that electron affinities for very small Si islands surrounded by an SiO_2 matrix are not severely modified from what is observed for thick single crystal, polycrystalline, or amorphous silicon layers themselves or in contact with an SiO_2 layer.[23] For this explanation, permanent electron trapping should not occur until very "deep" trapping sites in the SiO_2 are filled which can never communicate with the Si island conduction bands for any electric field condition. Electron capture into these very "deep" trapping sites would have to have a very low probability.

c) A final possibility which depends on the presence of hole injection into and conduction through these oxide films via the Si islands would be the annihilation or compensation by holes of the permanently trapped electron distribution in the oxide layer. A hole tunneling through the valence band states of the Si islands could have a finite probability of being trapped during the direct tunneling process. However, since there is no strong evidence for the motion of holes through these films, this possibility must be regarded with a certain amount of caution. Compensation of energetically deep trapped electrons in the SiO_2 phase of the oxide by mobile holes on the Si islands or by ionized donor type impurities in the Si phase seems a doubtful mechanism also from some of the experimental results. As discussed in section III-B (see Figs. 7 and 8), hysterisis in current-voltage characteristics of off-stoichiometric oxides due to permanent electron trapping in the SiO_2 phase could not be recovered over periods of days. If compensation by positive charges on Si islands (for example, by holes) was operative, this hysterisis should be partially or fully recovered by the motion of positive charges (holes) into the film via the Si islands to screen the local electric fields caused by the presence of the trapped electrons. If compensation of bulk trapped electrons by positive charge near the anode occurred at high fields and currents, then hysterisis effects in current-voltage characteristics should increase when the structure is returned to a low field condition for an extended period of time. In this case, the positive charges on the Si islands near the anode would be annihilated or compensated by mobile electrons tunneling in from the near-by contact and the full effect of the bulk trapped electrons on the internal electric field would be seen at low gate voltages.

This again was not seen experimentally since the hysterisis effects did not change in time under either grounded gate and substrate or floating conditions.

Any and all of the above possible explanations could account for the observed phenomena. Experiments designed to separate these possibilities are difficult if not impossible to do since the net effect on any of the measurable quantities (current, capacitance, electric field, trapped charge) using the techniques described here is very similar.

V. Application in EAROM Structures

The use of off-stoichiometric oxides (1-6% excess atomic Si) instead of stoichiometric SiO_2 in the intervening oxide layer of a DEIS EAROM structure results in extended cycling characteristics for these devices.[49] A schematic representation of a DEIS EAROM is shown in Fig. 2a. Electronic charge is transferred back and forth from the control gate to the floating gate by means of the enhanced injection and conduction of the modified DEIS (MDEIS) stack with its intervening off-stoichiometric oxide layer. This charge transfer is achieved at lower voltages and powers than a normal DEIS stack with an intervening layer of stoichiometric SiO_2 for pulse times $\gtrsim 500$ μsec.[49] The injector regions of the MDEIS or DEIS structures contain $\gtrsim 13\%$ excess atomic Si compared to the 1% to 6% excess atomic Si used in the intervening oxide layer. Writing or erasing (that is, putting negative charge on or taking it off the floating gate) is achieved by applying a negative or positive voltage to the control gate electrode, respectively.[14,15] An MDEIS stack with injectors can be written or erased at lower voltages as compared to a structure with just an off-stoichiometric oxide alone, as was discussed in section III-B (see Fig. 18). The threshold voltage V_T (that is, the gate voltage at which the channel of the device turns on[50]) of the DEIS or MDEIS EAROMs is used as a measure of the charge state (written or erased) of the floating gate electrode. Figure 27 shows the threshold voltage for the written V_{T_W} and erased V_{T_E} states as a function of the number of write/erase cycles for similar EAROM devices involving stoichiometric SiO_2 ($R_0 = 200$) and off-stoichiometric oxide ($R_0 = 50$) as intervening layers in DEIS stacks. The width of the voltage pulse for the write/erase operation was 500 μs. The threshold voltage shift between the written and the erased states $\Delta(\Delta V_T) = V_{T_W} - V_{T_E}$ (called the threshold voltage window) collapses in $\sim 10^3$ cycles from 6 V to ~ 2.5 V while a similar collapse takes $\sim 10^7$ cycles to occur in the structure with off-stoichiometric oxide. The improvement in the cycling characteristics of these devices results from the reduction of permanent electron trapping in the off-stoichiometric oxide layer. As some of the transferred electrons are permanently trapped, they create a repulsive internal electric field which lowers the electric

field near the injecting interface which in turn controls the current flow to and from the floating gate. As the interface field decreases, fewer electrons for the same write or erase voltage condition can be transferred and the threshold voltage shift between the written and erased states $\Delta(\Delta V_T)$ (which is a measure of the transferred charge) decreases. However, although an improvement of ≥ 4 orders of magnitude in the number of cycles is achieved, the threshold voltage window still collapses. This implies that permanent charge trapping in the oxide is still present. It is possible that a small percent of the electrons are still transported through the intervening oxide layer via the bottom of the conduction band of the SiO_2 phase giving rise to electron capture in deep traps (≥ 4 eV from the bottom of the SiO_2 conduction band) as was discussed in section E.

The permanent charge trapping efficiency is further reduced as the excess silicon content in the oxide is increased. This is illustrated in Fig. 28 where the cycling characteristics of devices incorporating off-stoichiometric oxides with different silicon content are shown. The least degradation for a given number of cycles is observed in the device involving an intervening oxide with the most excess Si content ($R_0 = 30$). The retention characteristics (that is, amount of a charge loss off the floating poly-Si gate electrode with time) at room temperature for a grounded control gate condition on the same series of devices as in Fig. 28 are shown in Fig. 29. The measurements were performed by charging the floating poly-Si gate of these devices with injected electrons to produce a threshold voltage of $\approx +8$ to $+9$ V from a virgin as-fabricated state and then the threshold voltage was monitored over at least a 24 hr. period. It was observed that as the silicon content is increased in the off-stoichiometric oxides (R_0 decreased) a slight degradation in the retention characteristics is observed. However, if the curves in Fig. 29 are corrected for read perturb effects (that is, the charge loss induced by the electric fields needed to determine the threshold voltage), all the data looks similar to the control structure with the stoichiometric intervening SiO_2 layer in the DEIS stack.

VI. Conclusions

The chemically vapor deposited off-stoichiometric oxides with 1% to 6% excess silicon have highly non-ohmic electrical characteristics that can be summarized as follows: i) good insulating qualities at low electric fields (≤ 1 MV/cm) where leakage current densities less than 10^{-11} A/cm^2 are observed, ii) current injection is achieved at lower electric fields than those required to induce equivalent current densities across stoichiometric oxide films, iii) a substantial reduction in the effects of permanent electron trapping normally present in stoichiometric SiO_2 is observed, iv) the temperature dependence of the I-V characteristics in

124

the range of $80^\circ K$ to $572^\circ K$ is similar to that observed for stoichiometric SiO_2. Therefore, the use of off-stoichiometric oxides instead of SiO_2 films in any kind of device requiring charge transfer across a thin oxide film for its operation will result in a better overall perform-ance. This has been demonstrated in particular for EAROM devices where an improvement of more than four orders of magnitude in the cycling characteristics is obtained with only slight, if any, degradation of the retention characteristics of these devices.

The excess silicon present in the off-stoichiometric oxide groups in clusters, probably ≤ 30 Å in diameter. These silicon islands behave as potential energy wells ≈ 3 eV to ≈ 4 eV deep from the bottom of the conduction band of the SiO_2. The dominant conduction mecha-nism appears to be controlled by the tunneling of electrons between silicon islands imbedded in the oxide. It is this difference in the electron transport mechanism (compared with normal stoichiometric SiO_2) that is believed to be one of the causes for the large reduction in permanent charge trapping efficiency observed in off-stoichiometric oxides. It is possible however, that a small component of the current is due to electron transport via the bottom of the conduction band of the SiO_2 phase. This current component will give origin to permanent electron capture in trapping centers believed to be located energetically in the bandgap of the SiO_2 phase, ≥ 4 eV from the bottom of the conduction band of the SiO_2.

Acknowledgements

References

1.   C. Falcony, D.J. DiMaria, D.W. Dong and K.M. DeMeyer, J. Appl. Phys. **53**, 43 (1982).

2.   R. Williams, Phys. Rev. **140**, A569 (1965).

3.   D.J. DiMaria, in **The Physics of SiO$_2$ and Its Interfaces**, ed. by S.T. Pantelides (Pergamon Press, New York, 1978), p. 160.

4.   E.H. Nicollian, C.N. Berglund, P.F. Schmidt, and J.M. Andrews, J. Appl. Phys. **42**, 5654 (1971).

5.   F.J. Feigl, D.R. Young, D.J. DiMaria, S. Lai, and J. Calise, J. Appl. Phys. **52**, 5665 (1981).

6.   C. Falcony-Guajardo, F.J. Feigl, and S.R. Butler, in **The Physics of MOS Insulators**, ed. by G. Lucovsky, S.T. Pantelides, and F.L. Galeener (Pergamon Press, New York, 1980) p. 127.

7.   D.D. Rathman, F.J. Feigl, S.R. Butler and W.B. Fowler, in **The Physics of MOS Insulators**, ed by G. Lucovsky, S.T. Pantelides, and F.L. Galeener (Pergamon Press, New York, 1980) p. 142.

8.   D.R. Young, J. Appl. Phys. **52**, 4090 (1981).

9.   S.K. Lai, D.R. Young, J.A. Calise, and F.J. Feigl, J. Appl. Phys. **52**, 5691 (1981).

10. C. Falcony, D.J. DiMaria, D.W. Dong, and S.K. Lai, *unpublished*.

11. D.W. Dong, E.A. Irene, and D.R. Young, J. Electrochem. Soc. **125**, 819 (1978).

12. T. Matsushita, T. Aoki, T. Otsu, H. Yamoto, H. Hayasi, M. Okayama, and Y. Kawana, IEEE Trans. Electron Devices **23**, 826 (1976).

13. D.J. DiMaria and D.W. Dong, J. Appl. Phys. **51**, 2722 (1980).

14. D.J. DiMaria, K.M. DeMeyer, and D.W. Dong, IEEE Elec. Dev. Lett. **EDL-1**, 179 (1980).

15. D.J. DiMaria, K.M. DeMeyer, C.M. Serrano and D.W. Dong, J. Appl. Phys. **52**, 4825 (1981).

16. D.W. Dong and P. Saunders, unpublished.

17. R.F. DeKeersmaecker and D.J. DiMaria, J. Appl. Phys. **51**, 1085 (1980).

18. A.M. Goodman, G. Harbeke, and E.F. Steigmeier, in **Physics of Semiconductors 1978**, ed. by B.L.H. Wilson (The Institute of Physics, Bristol and London, 1979), p. 805.

19. A. Hartstein, J.C. Tsang, D.J. DiMaria, and D.W. Dong, Appl. Phys. Lett. **36**, 836 (1980).

20. R.H. Alben, D. Wearie, J.E. Smith and M.H. Brodsky, Phys. Rev. B **11**, 2271 (1975).

21. S. Veprek, Z. Igbal, and F.A. Sarott, Philos. Mag. **B45**, 137 (1982).

22. Z. Iqbal and S. Veprek, J. Phys. C.: Solid St. Phys. **15**, 377 (1982).

23. G. Hollinger and F. Himpsel, *private communication*.

24. E.A. Irene and E. Tierney, *private communication*.

25. E.A. Irene, N.J. Chou, D.W. Dong and E. Tierney, Electrochem. Soc. **127**, 2518 (1980).

26. J.M. Gibson, in **Electron Microscopy 1980**, ed. by P. Brederoo and G. Boom (The Seventh European Congress on Electron Microscopy Foundation, Leiden, 1980), pp. 310-311, Vol. 1.

27. D.J. DiMaria and P.C. Arnett, IBM J. Res. Dev. **21**, 227 (1977).

28. R.J. Powell, J. Appl. Phys. **41**, 2424 (1970).

29. A.M. Goodman, Phys. Rev. **152**, 785 (1966).

30. R.M. Hill, Philos. Mag. **23**, 59 (1971).

31. D.J. DiMaria, R. Ghez, and D.W. Dong, J. Appl. Phys. **51**, 4830 (1980).

32. M. Lenzlinger and E.H. Snow, J. Appl. Phys. **40**, 278 (1969).

33. B. Abeles, P.Sheng, M.D. Couetts, and Y. Arie, in **Advances in Physics**, edited by D.H. Martin, B.R. Coles, and N.F. Mott (Taylor and Francis Ltd., London, 1975), pp. 407-461, Vol. 24.

34. A.M. Hartstein and D.R. Young, Appl. Phys. Lett. **38**, 631 (1981).

35. M. Av-Ron, M. Shatzkes, T.H. DiStefano, and R.H. Gdula, J. Appl. Phys. **52**, 2897 (1981).

36. D.J. DiMaria and D.W. Dong, J. Appl. Phys. **37**, 61 (1980).

37. J.R. Kirtley, T.N. Theis, J.C. Tsang, and D.J. DiMaria, submitted to Phys. Rev. B.

38. J.C. Knights, R.A. Street, and G. Lucovsky, J. of Non-Crystalline Solids **35 and 36**, 279 (1980).

39. D.J. DiMaria, in **The Physics of MOS Insulators**, ed. by G. Lucovsky, S.T. Pantelides, and F.L. Galeener, (Pergamon Press, New York, 1980), pp. 1-18.

40. Z.A. Weinberg, W.C. Johnson, and M.A. Lampert, Appl. Phys. Lett. **25**, 42 (1974).

41. A.S. Ginovker, V.A. Gritsenko, and S.P. Sinitsa, Phys. Status Solidi (a) **26**, 489 (1974).

42. P.C. Arnett and Z.A. Weinberg, IEEE Trans. Electron Dev. **ED-25**, 1014 (1978).

43. D.K. Schroder and M.H. White, IEEE Trans. Electron Dev. **ED-26**, 899 (1979).

44. J. Ni and E. Arnold, Appl. Phys. Lett. **39**, 554 (1981).

45. R. Car and S.T. Pantelides, *private communication*.

46. M.H. Brodsky, Solid St. Commun. **36**, 55 (1980).

47. R.M. Eisberg, **Fundamentals of Modern Physics**, (Wiley and Sons, New York, 1961), Chap. 8.

48. E. Merzbacher, **Quantum Mechanics**, (Wiley and Sons, New York, 1961), Chap. 7.

49.   D.J. DiMaria, D.W. Dong, C. Falcony, and S.D. Brorson, IEEE Electron Dev. Lett. **EDL-3**, 191 (1982).

50.   S.M. Sze, **Physics of Semiconductor Devices**, (Wiley-Interscience, New York, 1981), Chap. 8.
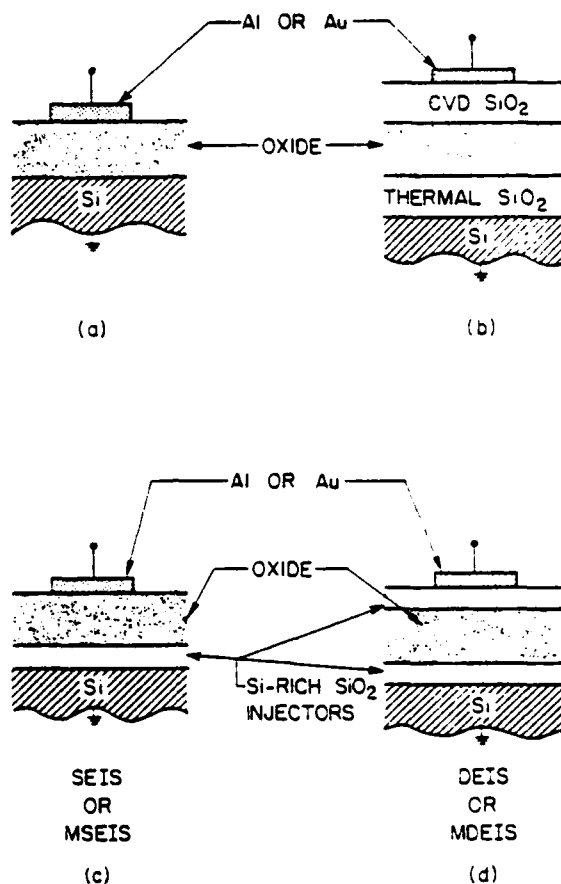
Fig. 1   Schematic representation of the types of capacitor structures used for many of the measurements discussed in this paper. (a) Simple MOS capacitor with either an SiO_2 or off-stoichiometric oxide sandwiched between the metal (Al or Au) electrode and the Si substrate (n or p-type <100>), (b) similar to (a) only with a thermal SiO_2/off-stoichiometric oxide/CVD SiO_2 dielectric stack used in the photodetrapping experiments, (c) SEIS (with an SiO_2 layer and one bottom Si-rich SiO_2 injector) and MSEIS (with an off-stoichiometric oxide and one bottom Si-rich SiO_2 injector) capacitor with metal (Al or Au) counter electrode and n-type degenerate ($\leq$ .001 Ωcm) Si substrate used for luminescence experiments, and (d) DEIS (an intervening SiO_2 and two Si-rich SiO_2 injectors) or MDEIS (an intervening off-stoichiometric oxide and two Si-rich SiO_2 injectors) capacitor similar to (a) except for the DEIS or MDEIS dielectric layer.
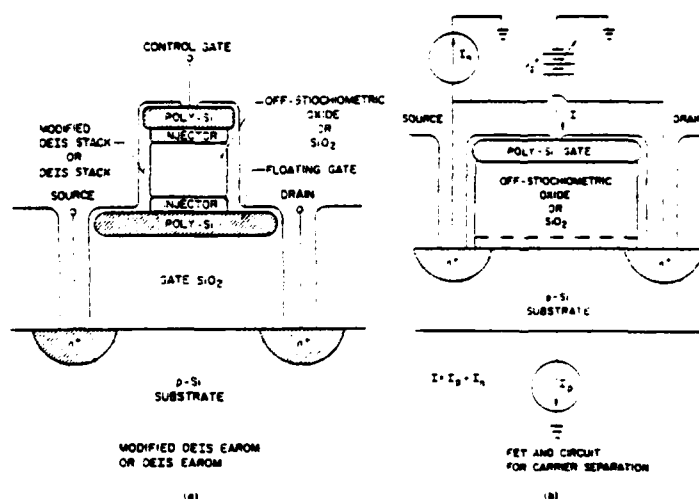
Fig. 2   (a) Schematic representation of a three port DEIS or MDEIS EAROM with an SiO$_2$ or off-stoichiometric oxide intervening layer in the DEIS or MDEIS stack, respectively. The poly-Si control (top) gate and floating (bottom) gate have areas of 1.3 x 10^{-6} cm^2 and 2.5 x 10^{-6} cm^2, respectively, at the masking level. The DEIS or MDEIS stack is incorporated inbetween the control and floating gates. A thermal oxide layer is grown from the Si substrate and used as a gate insulator inbetween the floating gate and the Si substrate. Not drawn to scale. (b) Schematic representation of the single poly-Si control gate FET and circuit used in the charge carrier separation experiments. The large control gate area was 1 x 10^{-4} cm^2 at the masking level and the gate oxide was formed from SiO$_2$, off-stoichiometric oxide, or a thin thermal SiO$_2$ layer inbetween a thicker off-stoichiometric oxide layer and the Si substrate (the thin SiO$_2$ boundary with the off-stoichiometric oxide is indicated by the dashed line in the drawing). Not drawn to scale.
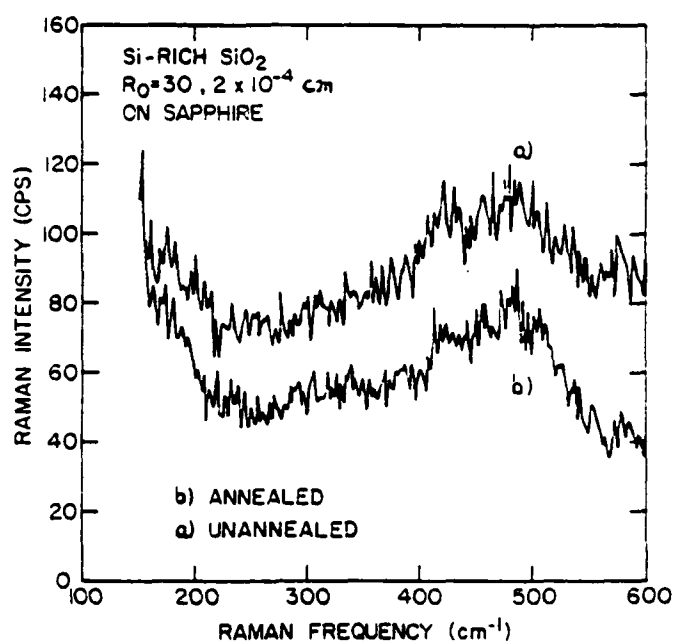
Fig. 3   Raman spectra at room temperature from 2 μm thick layers of R_0 = 30 oxide on a sapphire substrate. Curve (a) corresponds to an as-deposited sample while curve (b) corresponds to a sample annealed at 1000°C in N_2 for 30 min. Both samples show a broad Raman band near 500 cm^{-1} which is characteristic of Raman scattering from amorphous Si clusters in Si-rich SiO_2. The emission due to the sapphire substrate has been suppressed from this plot.
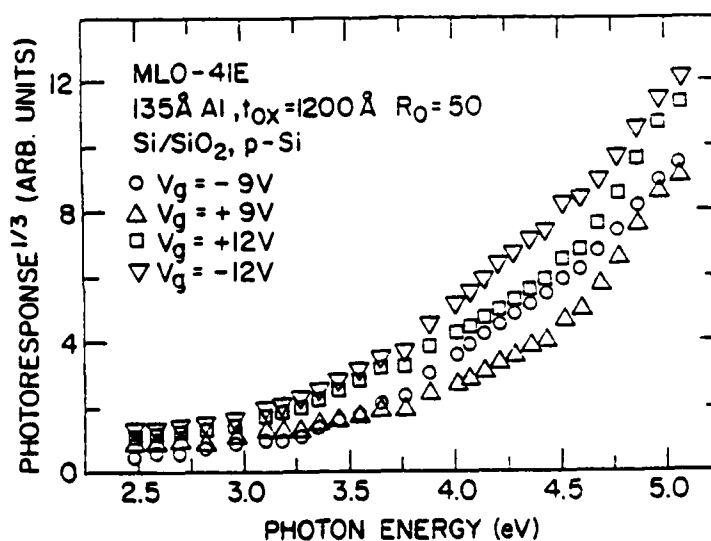
Fig. 4   Cube root of the d.c. photoresponse at room temperature as a function of the photon energy for a 1200 Å thick off-stoichiometric oxide (R_0 = 50) layer incorporated in an MOS structure. The metal electrode was a semi-transparent Al film (135 Å thick). The different curves correspond to applied gate voltages equal to ±9 V or ±12 V as indicated.
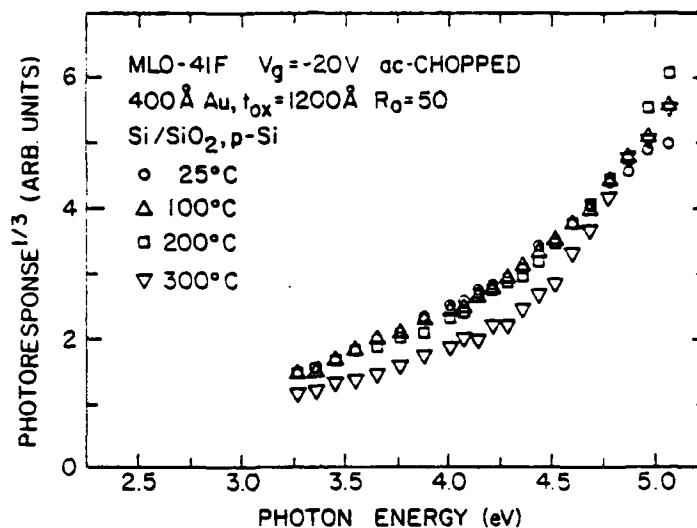
Fig. 5   Cube root of the a.c. chopped light photoresponse as a function of the photon energy for a 1200 Å thick off-stoichiometric oxide (R_o = 50) layer incorporated in an MOS structure. The metal electrode was a semi-transparent Au film (400 Å thick) and the gate bias for all data was -20 V. The different sets of data as indicated by the appropriate symbols correspond to temperatures of 25°C. 100°C, 200°C, and 300°C. Note the similarity between this figure and Fig 4.
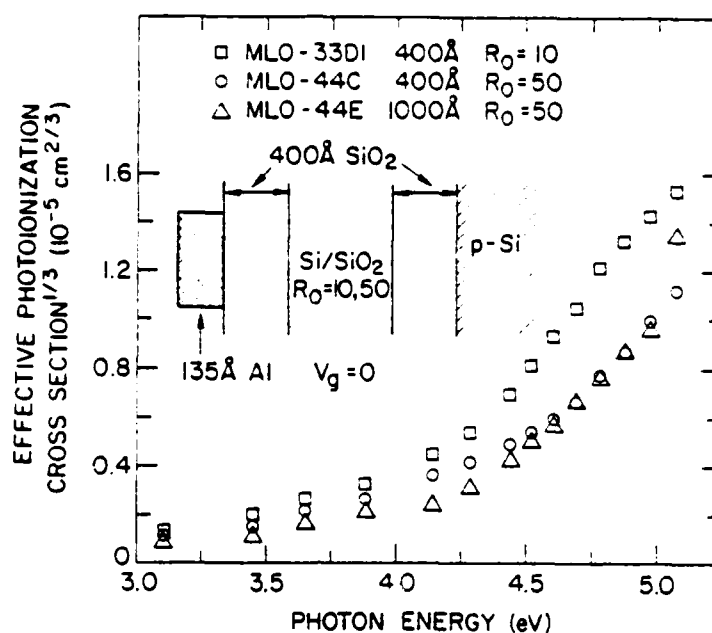
Fig. 6   Cube root of the effective photoionization cross section at room temperature as a function of photon energy for an off-stoichiometric oxide layer deposited between two 400 Å thick SiO₂ layers. Data for off-stoichiometric layers of various thicknesses (400 Å and 1000Å) and various excess Si content (R_o = 10 and 50) are shown. The curves show two thresholds at ~ 3 eV and at 4 eV. Note the similarity between this figure and Figs. 4 and 5.
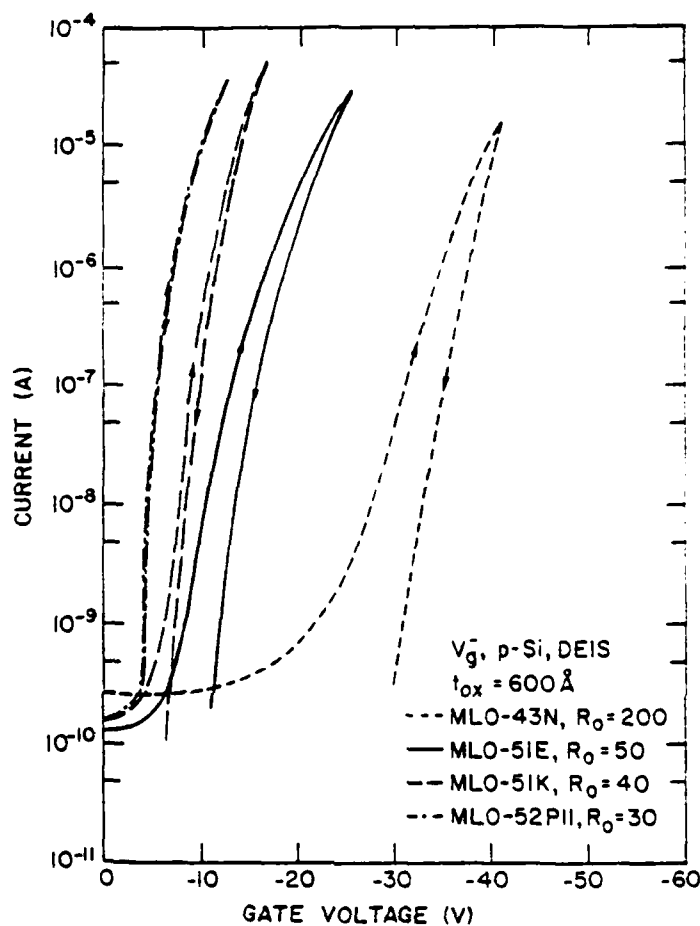
Fig. 7   Magnitude of the dark current at room temperature as a function of ramped gate
voltage (| dV_g/dt | = 0.5 V/sec) for SiO_2 and off-stoichiometric oxides with
1-2% (R_o = 50) to 5-6% (R_o = 30) excess silicon in a DEIS stack (see Fig. 1d).
The oxide layers were 600 Å thick, the Si-rich SiO_2 injectors (R_o = 3) were each
200 Å thick, and the metal gate electrode was Al.   A virgin capacitor was used
from each wafer.   As the excess silicon content increases in the oxide the I-V
curves shift towards lower gate voltages and the hysterisis associated with perma-
nent charge trapping diminishes.
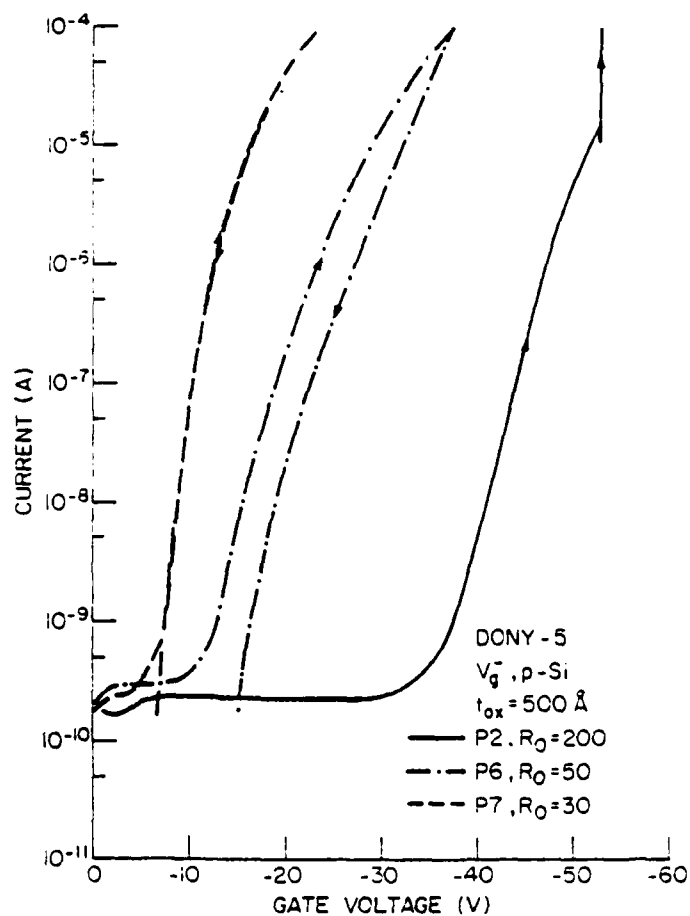
Fig. 8   Magnitude of the dark current at room temperature as a function of ramped gate voltage ($|$ dV_g/dt $|$ = 0.5 V/sec) for SiO_2 and off-stoichiometric oxides similar to those in Fig. 7 only with the oxide in an MOS configuration without the Si-rich SiO_2 injectors (see Fig. 1a). A virgin capacitor was used from each wafer. The behavior observed is similar to that observed in Fig. 7 except that the stoichiometric SiO_2 MOS capacitor could not be ramped to higher than 10^{-5} A without the oxide breaking down destructively.
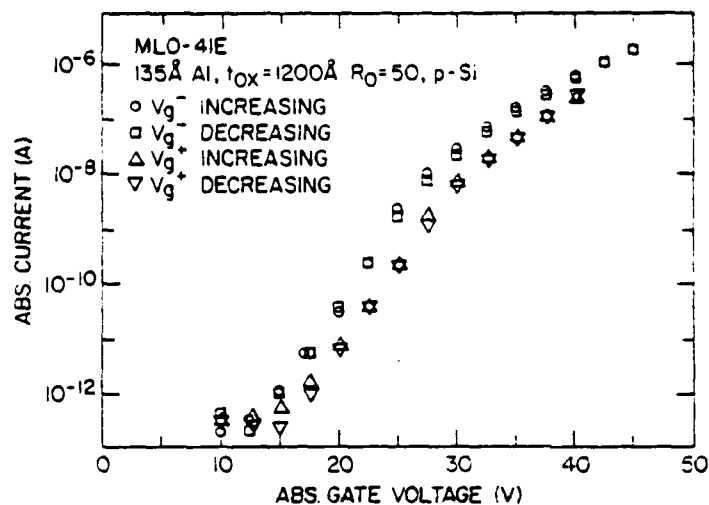
Fig. 9   Point-by-point magnitude of the dark current at room temperature as a function of the magnitude of the gate voltage for a 1200 Å thick $R_o = 50$ oxide in an MOS structure with an Al gate (see Fig. 1a). The leakage currents at low applied voltages ($\leq$ 15 V) are below the sensitivity of the picoammeter (10^{-12} A) (the actual current value is $< 10^{-12}$). The voltage required to obtain a given electron current across the device is approximately the same for both polarities.
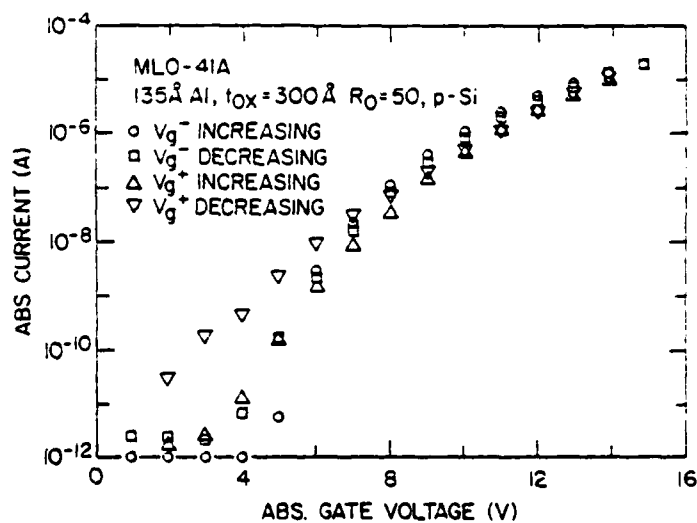
Fig. 10   Point-by-point magnitude of the dark current at room temperature as a function of the magnitude of the gate voltage for a 300 Å thick $R_o = 50$ oxide in an MOS structure with an Al gate (see Fig. 1a).
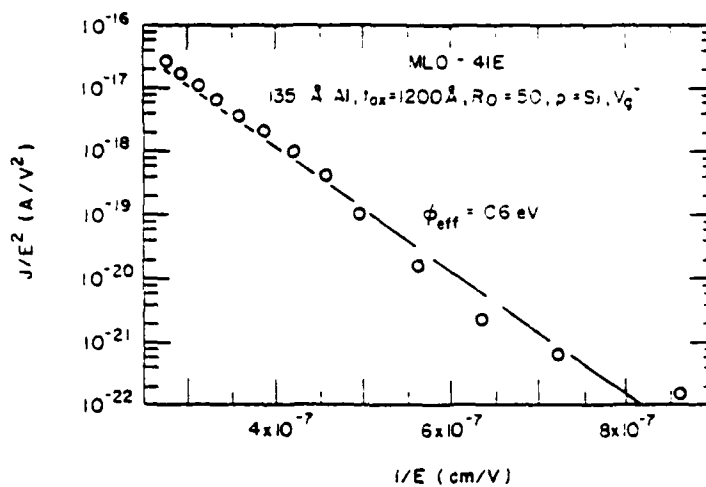
Fig. 11   Magnitude of the current density over the square of the magnitude of the average
electric field as a function of one over the magnitude of the average electric field
for negative gate voltage using the data shown in Fig. 9. The linear behavior of
the curve on a semilogarithmic plot suggests a tunneling conduction mechanism.
The full line represents a least-squares fit to the plotted $\ln (J/E^2)$ vs. $1/E$ data.
If the measured currents were due to Fowler-Nordheim tunneling, an effective
barrier height $\Phi_{eff} = 0.60$ eV (as deduced from the least-squares fit) would
correspond to this data.
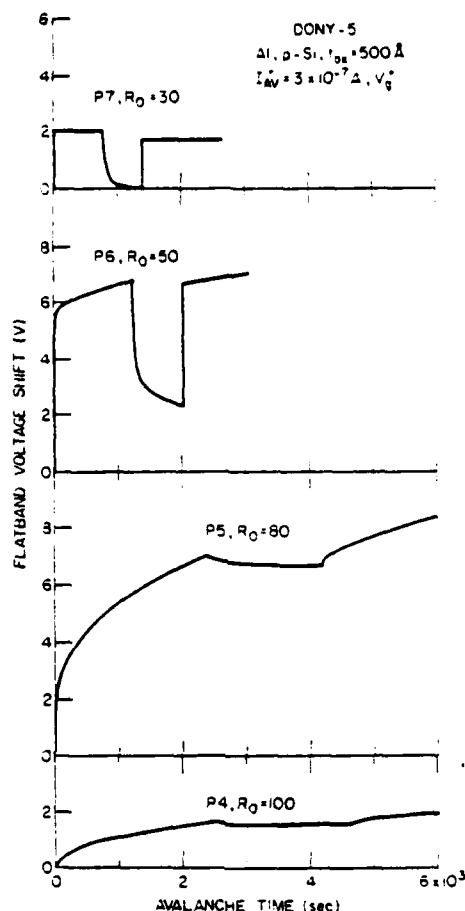
Fig. 12   Flatband voltage shift at room temperature as a function of avalanche injection time under a constant avalanche current of 3×10^{-7} A for various off-stoichiometric oxide layers on 0.2 Ωcm p-Si substrates incorporated into MOS capacitors with Al gate electrodes (see Fig. 1a). $R_o \geq 100$ corresponds to stoichiometric SiO_2 while $R_o = 30$-80 correspond to off-stoichiometric oxides with between 6% and ≤ 1% excess atomic Si incorporated into the film during deposition, respectively. The decreasing center portions of these curves correspond to where the avalanche current was shut-off and the metal control gate electrode was grounded.
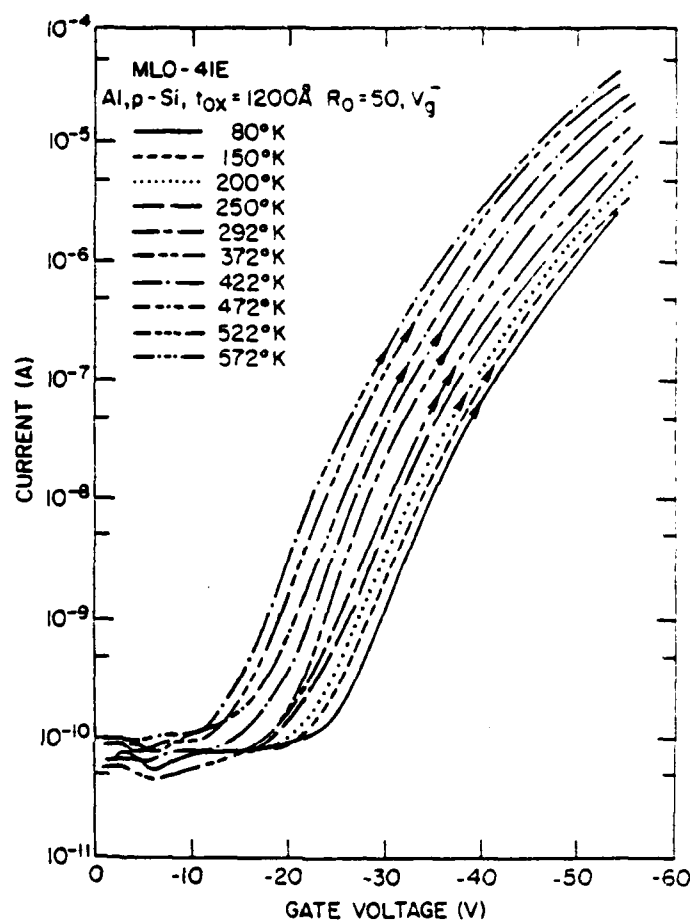
Fig. 13   Magnitude of the dark current as a function of ramped gate voltage ($|$ dV_g/dt $|$ $= 0.5$ V/sec) for various temperatures in the range from $80°$K to $572°$K on a 1200 Å off-stoichiometric oxide layer ($R_0 = 50$) incorporated into an MOS capacitor with an Al gate electrode (see Fig. 1a) under negative gate voltage bias. All measurements were performed on the same capacitor. The return trace which had a small amount of hystersis due to electron trapping was omitted for clarity. The I-V curves shift towards higher current values with increasing temperature. The total variation of the electric current values in this range of temperature is similar to that observed for stoichiometric SiO_2.
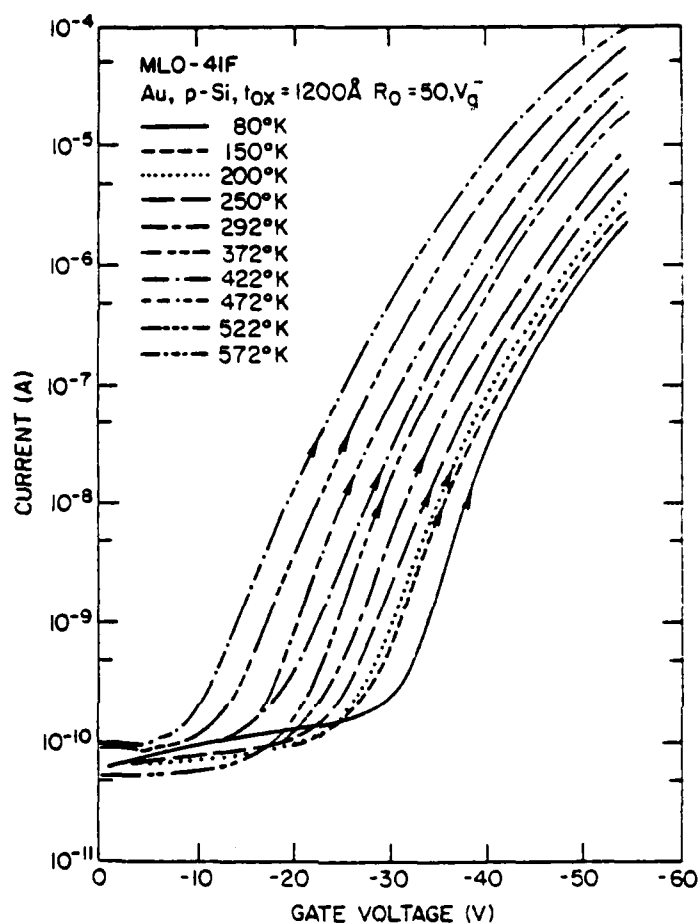
Fig. 14   Magnitude of the dark current as a function of ramped gate voltage ($| \, dV_g/dt \, |$ = 0.5 V/sec) for various temperatures in the range from 80°K to 572°K on a 1200 Å off-stoichiometric oxide layer (R_o = 50) incorporated into an MOS capacitor with an Au gate electrode (see Fig. 1a) under negative gate voltage bias. All measurements were performed on the same capacitor. The return trace which had a small amount of hystersis due to electron trapping was omitted for clarity.
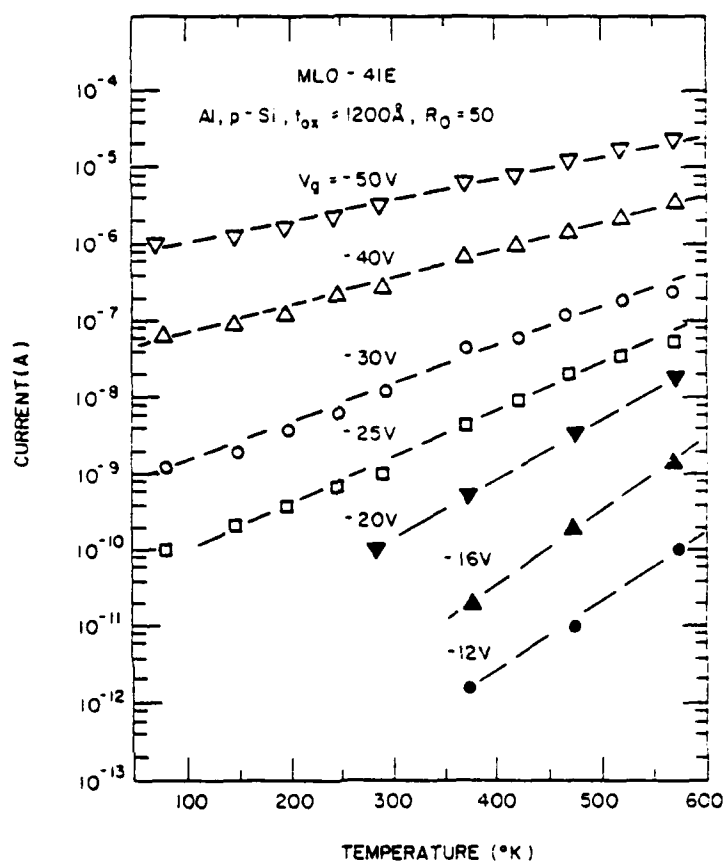
Fig. 15   Magnitude of the dark current as a function of temperature for several negative gate voltages on the same sample as in Fig. 13. Data for gate voltages between -12 V and -20 V were obtained from point by point measurements, while data for gate voltages from -25 V to -50 V were obtained from ramp I-V measurements. This data shows that the temperature dependence decreases with increasing electric field.
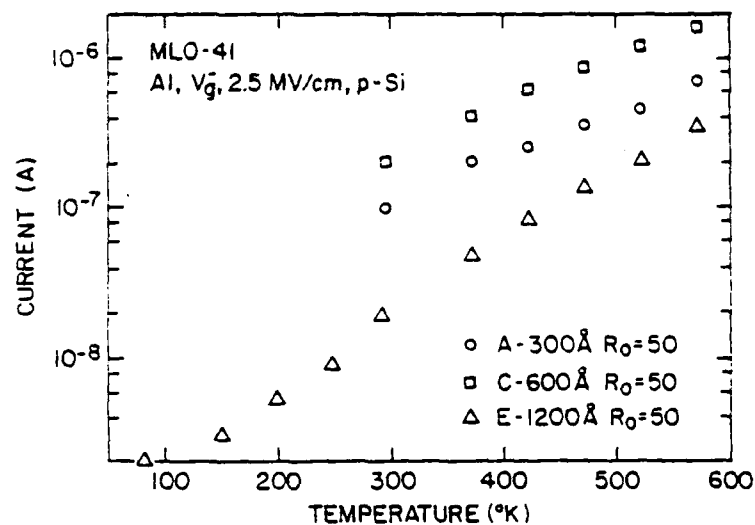
Fig. 16   Current as a function of temperature from 80°K to 572°K under negative gate voltage bias with an average electric field magnitude of 2.5 MV/cm across the off-stoichiometric oxide for R_o = 50 material of varying thickness (300 - 1200 Å) incorporated into MOS capacitors with Al gate electrodes (see Fig. 1a). Note the weak temperature dependence and similarity of the data for the various oxide thicknesses.
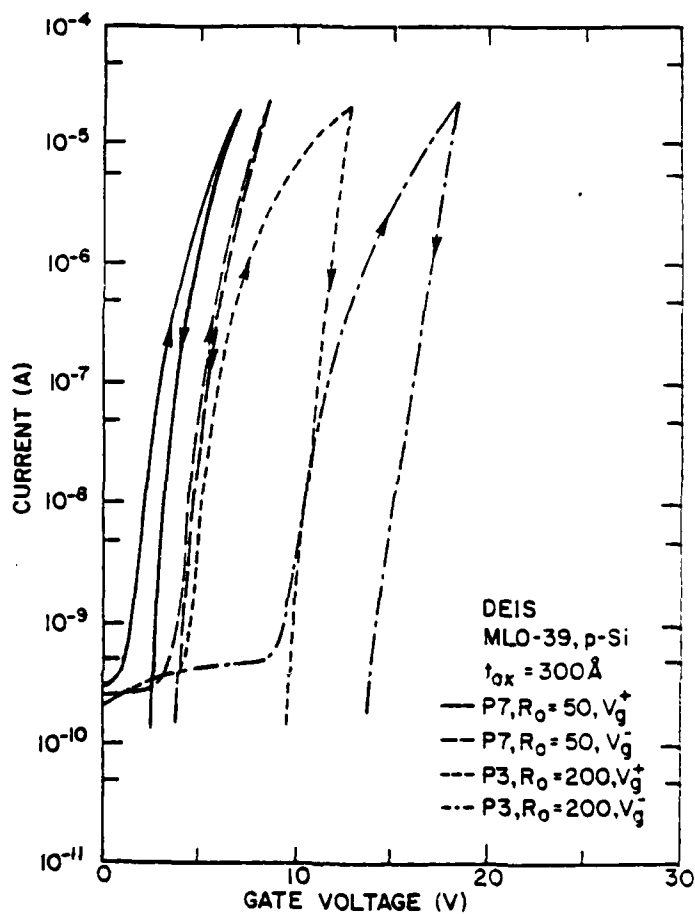
Fig. 17   Magnitude of the dark current at room temperature as a function of the magnitude of the ramped gate voltage (| dV_g/dt | = 0.5 V/sec) for DEIS structures (see Fig. 1d) with either an SiO_2 (R_o = 200) or an off-stoichiometric (R_o = 50) intervening oxide layer which was 300 Å thick. The Si-rich SiO_2 injectors (R_o = 3) were each 200 Å thick and the gate electrode was Al. Both polarities are shown with a virgin capacitor ramped for each polarity.
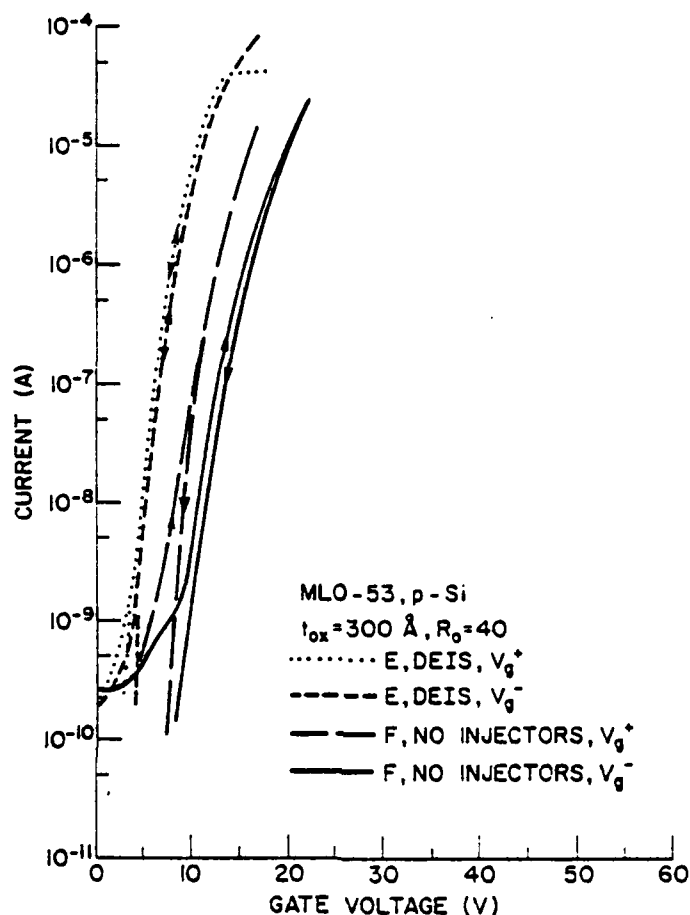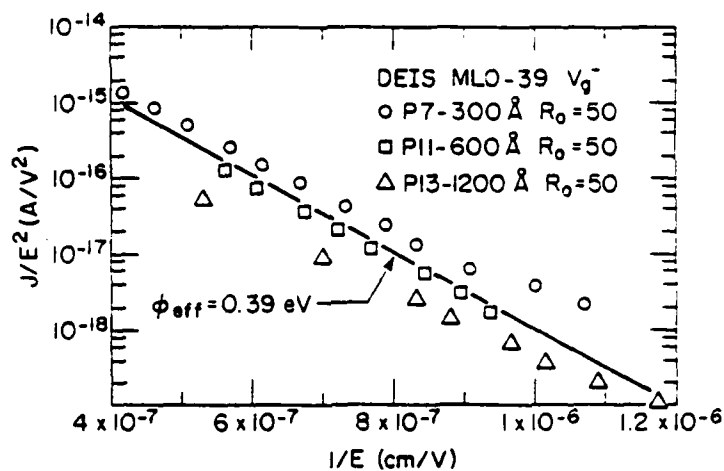
Fig. 18   Magnitude of the dark current at room temperature as a function of the magni-
tude of the ramped gate voltage ($| dV_g/dt | = 0.5$ V/sec) for capacitors using
off-stoichiometric oxides ($R_o = 40$) with (see Fig. 1d) and without (see Fig. 1a)
Si-rich SiO_2 ($R_o = 3$) injectors. The oxide layers were 300 Å thick, the Si-rich
SiO_2 injectors were each 200 Å thick, and the gate electrode was Al. Both
polarities are shown with a virgin capacitor ramped for each polarity.

145

Fig. 19   Magnitude of the current density at room temperature over the square of the
magnitude of the average electric field as a function of one over the magnitude of
the average electric field for negative gate voltage for DEIS capacitors (see Fig.
1d) with off-stoichiometric (R_o = 50) intervening oxide layers of various thick-
ness (300 - 1200 Å). The Si-rich SiO_2 injectors were each 200 Å thick and the
gate electrode was Al. The full line represents a least-squares fit to all plotted ℓn
(J/E^2) vs. $1/E$ data. If the measured currents were due to Fowler-Nordheim
tunneling, an effective barrier height of Φ_{eff} = 0.39 eV (as deduced from the
least-squares fit) would correspond to this data.

Fig. 20   Magnitude of the current density at room temperature over the square of the
           magnitude of the average electric field as a function of one over the magnitude of
           the average electric field for positive gate voltage on the same series of DEIS
           capacitor wafers used for the data in Fig. 18. Similar to Fig. 18, the full line
           represents a least square fit to the data assuming the currents measured in the
           external circuit are limited by or behave similarly to a Fowler-Nordheim tunneling
           mechanism. The effective barrier height deduced from fitting all the data is Φ_{eff}
           = 0.36 eV.

Fig. 21   Schematic energy band representation of an MOS capacitor system [n-type
           degenerate (.001 Ωcm) Si/Si-rich SiO_2/SiO_2/Au] used for the luminescence
           experiments (Fig. 21a). In certain experiments the SiO_2 layer was replaced by an
           off-stoichiometric oxide ($\approx$ 1-6% excess atomic Si) represented in Fig. 21b which
           is thought to perturb the SiO_2 bands because of the presence of small Si inclu-
           sions similar to the Si-rich SiO_2 injector layers which have $\gtrsim$ 13% excess atomic
           Si.

147

Fig. 22   Light output (luminescence) at room temperature as a function of energy for the MOS capacitor structures (.01 cm^2 gate area) with SiO_2 and off-stoichiometric oxides as shown schematically in Fig. 1c and Fig. 21. Injector current was held constant at an average value of 5×10^{-6} A by circuits which continually adjusted the positive gate voltage bias. Wavelength was swept from high to low photon energy (increasing wavelength) These data were corrected for the spectral response of the mirror optics, monochromatic grating and the photo tube used. The Si-rich SiO_2 injector, oxide layer, and Au were approximately 200 Å, 500 Å and 250 Å in thickness, respectively. The • and ▲ symbols correspond to the luminescence data from structures with stoichiometric SiO_2 and with R_0 = 40 oxides (2-3% excess atomic Si), respectively. The Si-rich SiO_2 injector had ≥ 13% excess atomic Si (R_0 = 3).
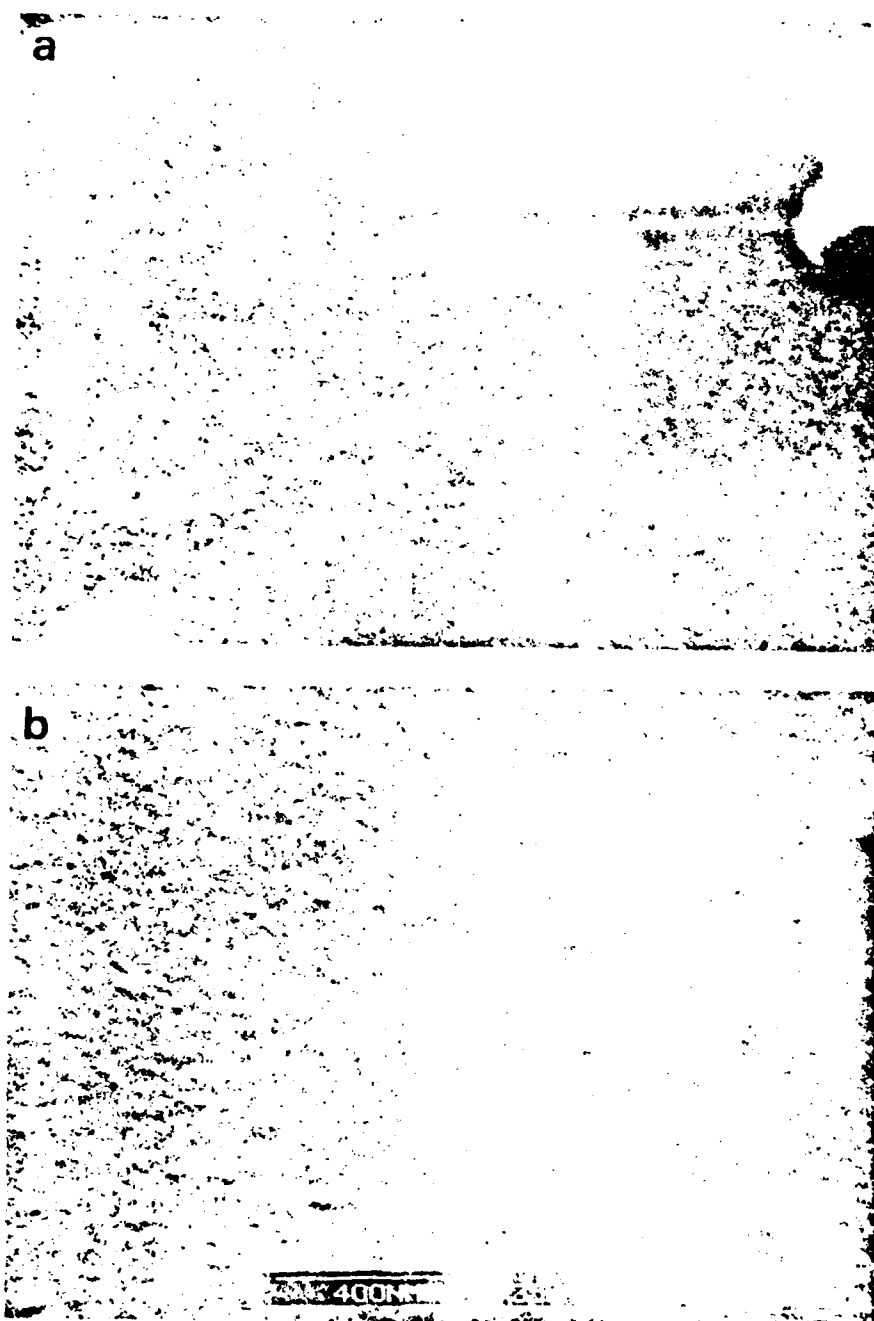
Fig. 23   SEMs using a 35 KV electron beam with the MOS wafer tilted at a 45° angle to the incident electron-beam for (a) 2 Ωcm n-Si <100>, 500 Å CVD SiO_2 (R_0 = 200)/300 Å Au structure, and (b) 2 Ωcm n-Si <100> 500 Å CVD off-stoichiometric SiO_2 (R_0 = 30)/300 Å Au structure.
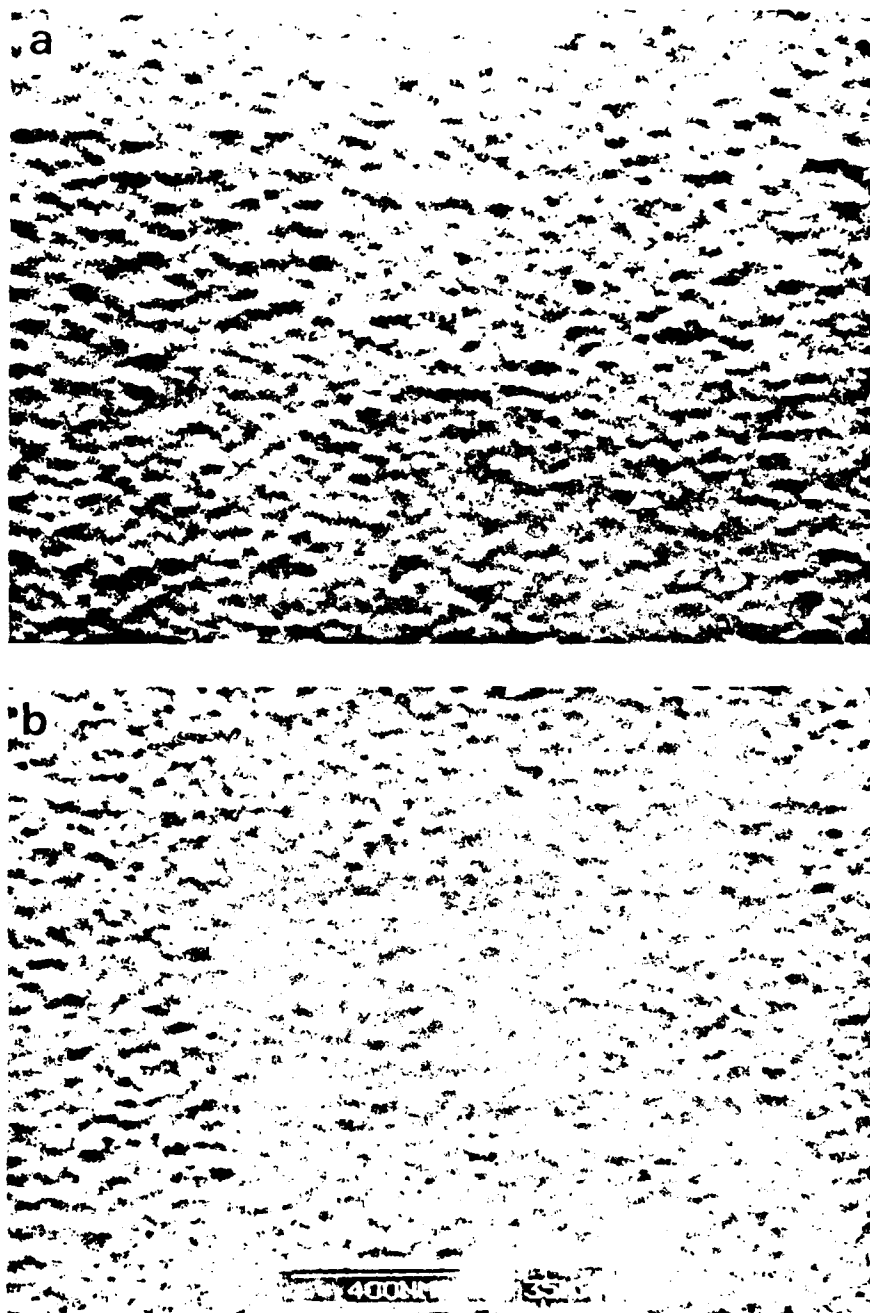
Fig. 24   SEMs using a 35 KV electron beam with the MOS wafer tilted at a 45° angle to
the incident electron-beam for (a) 2 Ωcm n-Si <100>/200 Å CVD Si-rich SiO_2
injector (R_o = 3)/500 Å CVD SiO_2 (R_o = 200)/300 Å Au structure and (b) 2
Ωcm n-Si <100>/200 Å CVD Si-rich SiO_2 injector (R_o = 3)/500 Å CVD
off-stoichiometric SiO_2 (R_o = 30)/300 Å Au structure.
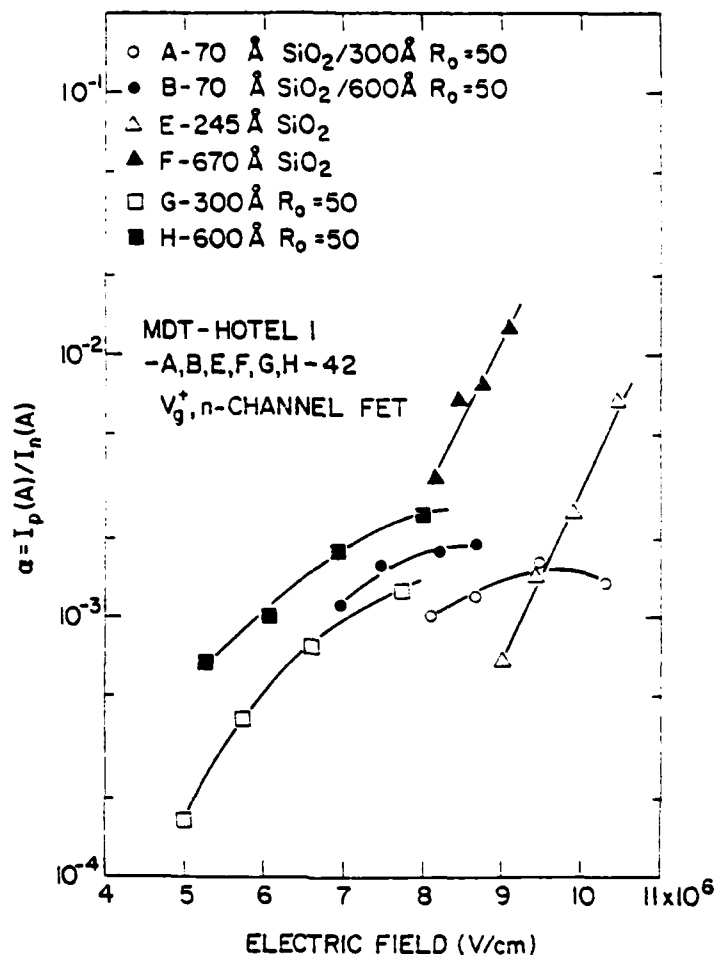
Fig. 25   The ratio of 0.5 Ωcm p-type Si substrate current to n^+ diffusion current $\alpha = I_p/I_n$

for the n-channel FET devices (MDT-HOTEL 1-A.B.E.F.G.H-42) as a function

of the average electric field across the oxide layer.   The device geometry and

circuit configuration for the carrier separation measurements are shown in Fig.

2b.   All measurements were done in the dark at room temperature.   Wafers G and

H had 300 Å and 600 Å of $R_0 = 50$ (1-2% excess atomic Si) off-stoichiometric

oxides for the gate insulator: wafers A and B had 70 Å of thermal $SiO_2/300$ or

600 Å of $R_0 = 50$ oxide stacks for the gate insulator: and wafers E and F had

245 Å and 670 Å of thermal SiO_2 for the gate insulator.   The control gate area

was 1×10^{-4} cm^2, at the masking level.   This data shows that little, if any, hole

current was observed over a wide range of average electric fields for positive

voltages.
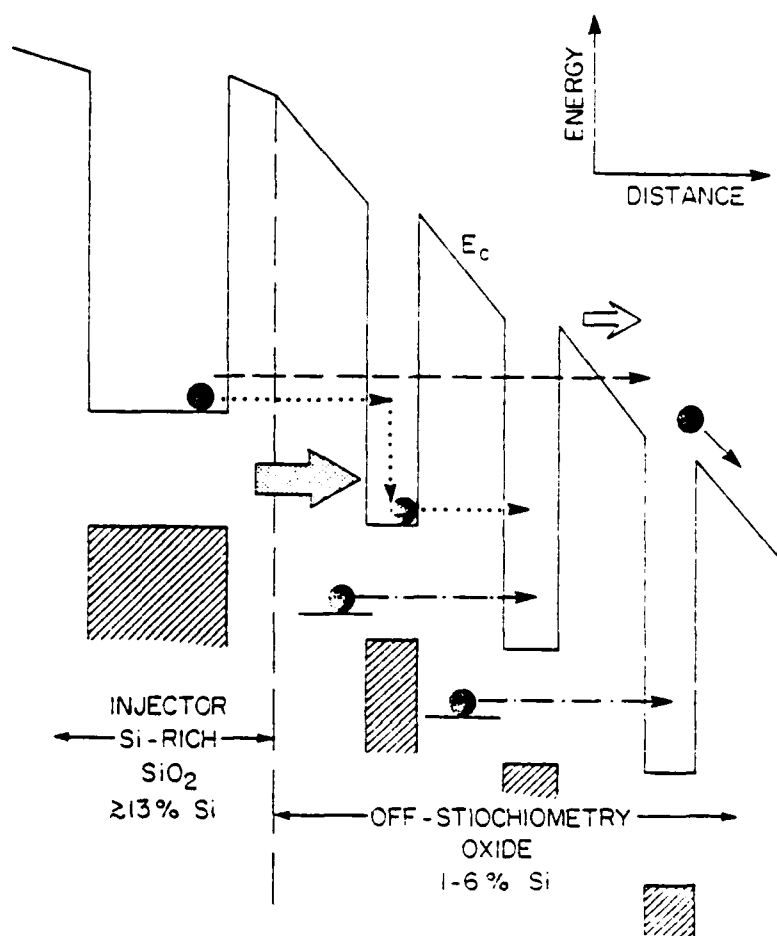
Fig. 26 Schematic energy band representation of the Si-rich SiO$_2$ injector off-stoichiometric oxide interfacial region. The dashed arrow shows the electron path for the localized field-enhanced Fowler-Nordheim tunneling into the SiO$_2$ conduction band for the SiO$_2$ portion of the off-stoichiometric oxide layer. The dotted arrow shows the direct tunneling path between the small Si regions of the off-stoichiometric oxide which is believed to dominate the conduction mechanism. Dot-dash arrow shows the direct tunneling path of trapped electrons on energetically "deep" SiO$_2$ sites which can tunnel to the Si regions in the off-stoichiometric oxide under favorable electric field conditions.
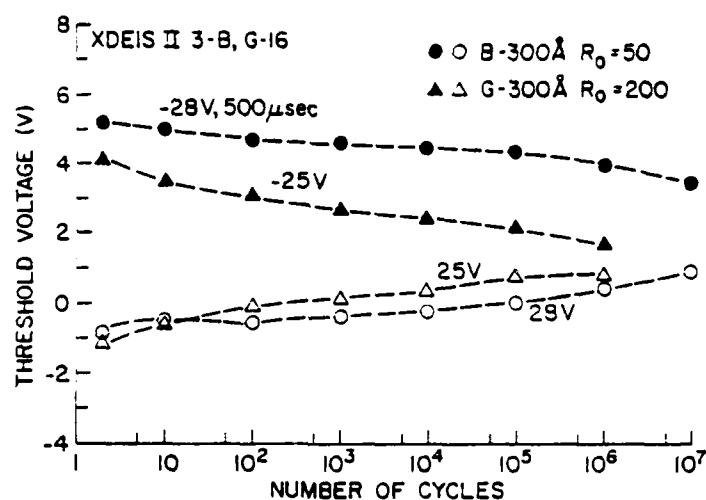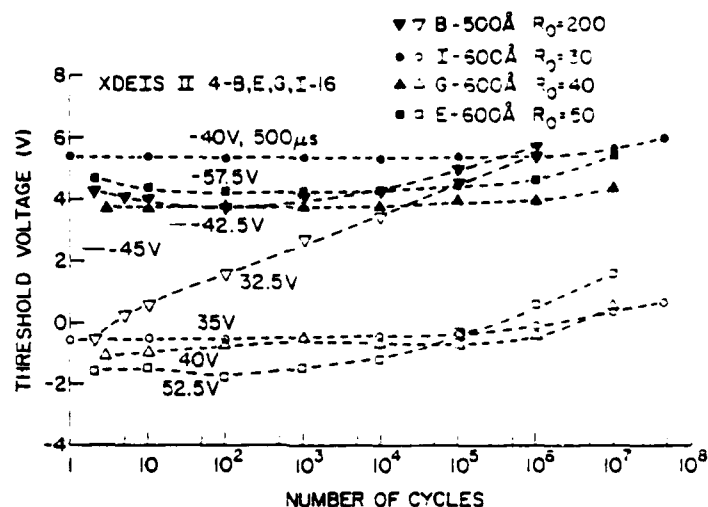
Fig. 27   Cycling characteristics for DEIS or MDEIS EAROM devices (XDEISII 3-B,G-16) involving a 300 Å thick intervening stoichiometric oxide (R_o = 200) or 300 Å thick intervening off-stoichiometric oxide (R_o = 50) in the DEIS stack. The write/erase operation of the devices was performed with square voltage pulses 500 μs in duration at room temperature. The write/erase voltage amplitudes are indicated in the plot for each case. The device geometry is depicted in Fig. 2a. The gate oxide is 650 Å thick. The control gate and floating gate had areas of 1.3 x 10^{-6} and 2.5 x 10^{-6} cm^2, respectively, at the masking level. The off-stoichiometric oxide sample shows a slower collapse of the threshold voltage window than that observed on the sample with stoichiometric oxide.

Fig. 28   Cycling characteristics for DEIS or MDEIS EAROM devices (XDEISII 4-B,E,G,I-16) involving 500 Å thick intervening stoichiometric oxide (R_0 = 200) or 600 Å thick intervening off-stoichiometric oxide layers (R_0 = 30, 40, 50) in the DEIS stack. The write/erase operation of the devices was performed with square voltage pulses 500 μs in duration at room temperature. The write/erase voltage amplitudes are indicated in the plot for each case. The device geometry is depicted in Fig. 2a. The gate oxide is 850 Å thick. The control gate and floating gate had areas of 1.3 x 10^{-6} and 2.5 x 10^{-6} cm^2, respectively, at the masking level. Better cycling characteristics are observed as the excess silicon content in the oxide is increased.
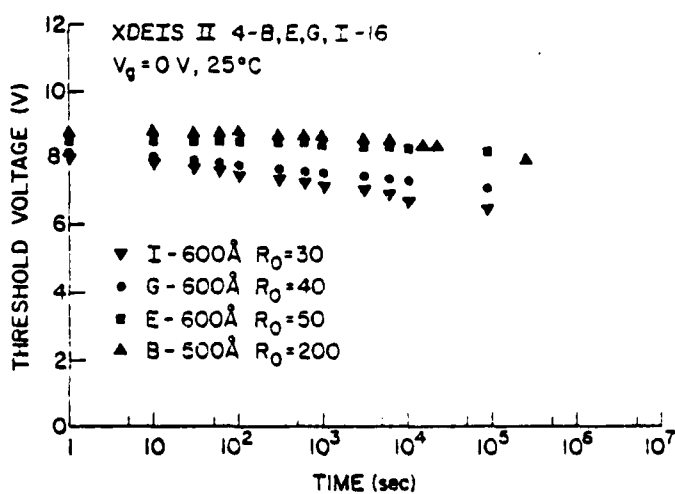
Fig. 29   Charge retention characteristics (threshold voltage as a function of time) at room temperature for a grounded control gate condition on the same series of devices as in Fig. 28. The initial charge state of the floating gate electrode was set at V_1 ≈ +8 to +9 V (stored electrons).

154

**Silicon-Rich SiO$_2$ and Thermal SiO$_2$ Dual Dielectric
for Yield Improvement and High Capacitance** [*]

S.K. Lai, D.J. DiMaria and F.F. Fang

I.B.M. Thomas J. Watson Research Center
Yorktown Heights, New York 10598

Technical Assistance by F.L. Pesavento and Silicon Processing Facility

Abstract:

The use of silicon-rich SiO$_2$ and thermal SiO$_2$ dual dielectric in memory capacitors and FETs was investigated. It was shown that the silicon-rich layer was conductive and introduced only a small decrease in the series capacitance of the dual dielectric. Consequently, the capacitance of the dual dielectric was close to that of the thermal oxide only. The response time of the silicon rich layer was measured by using FET response time and was shown to be in the nanosecond range. With this fast response time, it is possible to use the dual dielectric in memory and logic circuits. Another advantage of the dual dielectric is the very high yield due to the field screening of the silicon-rich layer to any non-uniformities in the thermal oxide or at the SiO$_2$-contact interface. This dual dielectric has the promise of high yield and high capacitance for future VLSI circuits.

I. Introduction

In VLSI circuits, a high-dielectric-constant gate-material is very desirable. It can increase storage capacitance in 1-Device dynamic memory cells[1] to increase signal levels and reduce impact of alpha particles[2], increase transconductance of FETs, and reduce short channel effects[3]. Insulators like silicon nitride[4] and tantalum pentoxide[5] have been proposed and used to give higher dielectric constant compared to that of thermal oxide. However, the use of such insulators is still experimental, and their compatibility with standard polysilicon gate processing and long term reliability are questionable.

Recently, it was demonstrated that silicon-rich silicon dioxide (deposited by CVD with approximately 13 percent excess silicon)[6] could be used on top of thermal silicon dioxide for enhanced electron injection into the oxide[7-9]. Furthermore, silicon-rich oxide has a high dielectric constant (approximately 7.5 in the above composition) and such a composite structure has been demonstrated to have well controlled high field conduction properties[9]. In the low field region (less than 5 MV/cm) where there was no significant injection of electrons, such a dual dielectric gave higher capacitance compared to pure oxide of the same thickness and had very few low field breakdowns, giving high yields[9]. It will be shown in this article that under suitable conditions, the response time of stored charge in the silicon-rich layer can be in the nanosecond range, fast enough to be used in most circuits. From a processing point of view, the use of silicon-rich oxide is only a minor perturbation in normal

polysilicon gate processing[9]. The process of deposition and patterning of silicon-rich oxide is compatible with all existing technologies and it can be processed at high temperature if it is protected by polysilicon from direct exposure to oxygen. Some field-effect transistors (FETs) have been fabricated to investigate the use of the composite insulators. The results of the experiment will be reported here.

II. Capacitance and transconductance

A total of six wafers were processed with different thermal oxide and silicon-rich oxide thicknesses. The mask set was a test site with discrete FETs of different dimensions, together with capacitors and other test structures. A standard single level polysilicon gate process was used with no ion implantation. The additional steps needed for the present experiment were the deposition of silicon-rich oxide before polysilicon and the reactive ion etching of the silicon-rich layer using the etched polysilicon gate as a mask. The silicon-rich oxide layer can be considered an extension of the polysilicon layer. Table I shows the oxide thicknesses for wafers A to F. The three wafers with silicon-rich oxide all have 10 nm thermal oxides. Figure 1 shows the 1 MHz capacitance (small signal: 15 mV rms) curves for square capacitors 10 mils on a side. It can be seen that wafers with silicon-rich oxide all have higher capacitance compared to the one with 25 nm thermal oxide. A 40 nm silicon-rich oxide layer was equivalent to approximately 6 nm of thermal oxide. The effective dielectric constant of the silicon-rich oxide was determined to be

between 20 to 30. This was higher than what was reported earlier[7]. The high dielectric constant is due to the conductance of and charge storage in the silicon-rich layer which was believed to be more conductive in the present set of samples. The higher conductivity was the result of either higher silicon content or possibly doping of the silicon-rich layer by phosphorous dopants from the gate region. The conductivity of silicon-rich oxide is non-linear and is a strong function of electric field[6]. For any finite voltage drop across the layer, current will start to flow; however, the external circuit current is limited by the oxide barrier. The conductivity of the Si-rich SiO_2 and charge build-up in this layer (particularly at its interface with the underlying SiO_2 layer) decreases the field in the bulk of Si-rich SiO_2. In the steady state, there is a small voltage drop across the layer to support the charge redistribution in the layer. The additional charge stored gives a higher measured capacitance compared to the case if silicon-rich oxide is insulating and has a dielectric constant of about 7.5. As a result, more charge is stored in the device. The change in capacitance under bias has been reported before[7,8] by ramp I-V capacitance measurement. Similar results were obtained recently for a 1 MHz capacitance measurement[10]. The high capacitance also gave high transconductance for the FETs, as could be seen in Figure 2. There was a larger variation of transconductance from device to device in wafer D and E due to limitations of the atmospheric pressure CVD system which gave approximately 10 percent variations in thickness and silicon content for the silicon-rich oxide

layer. This could easily be improved by better deposition techniques like plasma or low pressure CVD processes.

The most important difference between the use of silicon-rich oxide and silicon nitride which has been proposed is due to the conductivity of the silicon-rich oxide. The conductivity of the silicon-rich oxide eliminates the effects of permanent charge trapping which can be large for silicon nitride. Therefore, it is expected that the electron trapping of the silicon-rich oxide and thermal oxide composite is simply that of the thermal oxide only. In conclusion, it is possible to obtain high capacitance even with relatively thick layers of silicon-rich oxide.

III. Response time

One main concern about the use of silicon-rich oxide is the charge stored in the layer itself. The charging and discharging of the layer takes a finite length of time. The response time of the silicon-rich oxide is studied by examining the change of source to drain current in response to a change in gate to source voltage. The results for wafer A and D are shown in Figure 3. Within the accuracy of this measurement, there is little difference in the response time between the two cases. Similar results were obtained for wafers E and F. Even wafer F which had 40 nm of silicon-rich oxide did not show any additional delay due to the relatively thick silicon-rich layer. This very fast response time of the silicon-rich layer is due to the fact that it is conduc-

tive. In response to pulses, the silicon-rich oxide is put into a very conductive regime, giving the fast charging of the layer. Since the conductivity is almost symmetrical under positive and negative field, the charging and discharging was very fast and this was shown when the switch off response of the same devices was examined. It is difficult to examine the response time of the silicon-rich oxide by measuring single FETs alone. For more accurate measurements, actual circuits or ring oscillators must be fabricated to measure the limitation of the response time of silicon-rich oxide.

IV. Breakdown and Yield

The breakdown properties of the composite layer have been reported previously[9] and similar results were observed in the present experiment. Traditionally, breakdown testing was performed with the gate under negative bias so that the P type substrate was in accumulation. This is actually not the relevant bias condition for FETs which are biased with positive voltages. But if the storage capacitor in the 1-Device memory cell is formed by a N^+ diffusion, then the negative bias condition is the appropriate bias condition. Under negative bias, the top gate with silicon-rich oxide is the injecting contact and the breakdown improvement under such conditions is well documented[9]. This is due to the screening effect of the silicon-rich oxide on injected charge from any inhomogeneous high field spots which are responsible for low field breakdowns in thermal SiO_2. The drawback is the field at which there is significant carrier injection into the oxide is lowered from about

8 MV/cm to 5 MV/cm[7] if the thickness of the silicon-rich oxide layer is ignored. In effect, one is trading low field breakdowns and scattered distribution for a case where the maximum breakdown field is lowered but the distribution is tight and controllable. The results of such measurements on 10 mil square capacitors are shown in Figure 4. The ramp rate was set at 0.5 MV/cm-sec and the current limit was approximately 0.16 Amp/cm^2. The breakdown events are actual destructive breakdown of the oxide.

The problem of breakdown measurement under positive bias is that unless there is a source of electrons to supply carriers to the surface, the depletion region in the silicon limits the current supplied. Any increase in voltage goes into the depletion region so that the oxide voltage is not increased and thus there is no breakdown. In the present case, a FET was used for the measurement. The source and drain were tied to the substrate to give the surface channel a sufficient supply of carriers. In this case, the injecting contact was the silicon substrate and the silicon-rich oxide did not screen the field at inhomogeneities directly. Experimentally (Figure 5), it was observed that there was some improvement, even for this case. The ramp rate was 0.5 MV/cm-sec and the current limit was approximately 1 Amp/cm^2. Again, the breakdown events were actual destructive breakdowns. The exact mechanism for the improvement is still being studied. In conclusion, the breakdown and yield of the Si-rich oxide and thermal oxide composite is superior to that of the thermal oxide only, with the maximum breakdown field lowered in the

dual dielectric. The impact in the future when very thin gate dielectric are used in submicron circuits is obvious.

V. Work function and threshold voltage

One interesting question raised by the use of silicon-rich oxide is the effect of this material on the effective work function difference between the polysilicon gate and the silicon substrate and how this would affect the threshold voltage. To first order, silicon-rich oxide is composed of silicon islands embedded in a silicon dioxide matrix and in this composition range, it is basically like an oxide. The layer thus would not affect the work function difference. The threshold voltage should be close to that of straight thermal oxide only. This was demonstrated when the threshold voltages were measured. The threshold voltage for FETs in wafer D with 10 nm of silicon-rich oxide was only 50 mV higher than that of wafer A. Furthermore, the scatter in the threshold voltage in the two wafers were similar and was basically limited by the scatter in the thickness of the underlying thermal oxide. A silicon-rich oxide layer thus would not affect any technique that is used to control the threshold voltage, and yet results in an improvement in breakdown and yield. Furthermore, it is even possible to use other gate materials on top of the silicon-rich oxide that might otherwise damage the gate oxide.

VI. Discussion

In effect, one can look at silicon-rich oxide as a new kind of gate contact material. Electrically, in capacitance and transconductance, the silicon-rich oxide layer behaves like a very thin oxide layer. Physically, one has a much thicker dielectric layer for the same capacity per unit area. In the present experiment, the silicon-rich oxide is patterned by the polysilicon layer and is simply an extension of the polysilicon layer. The two layers can be deposited one after another in the same apparatus.

There is one important variable which is not optimized in the present case: the amount of silicon in the silicon-rich oxide. In the present CVD system, the amount of excess silicon is limited to approximately 13 atomic percent. The charge stored and response time will further be improved with higher silicon content. There is of course a limit to the improvement because in the extreme, this layer becomes equal to polysilicon. It is important to determine *the point at which one still* has significant breakdown improvement due to field screening together with the advantage of large charge storage and fast response time.

In the long term, the most important question is the reliability of the *silicon-rich oxide and thermal oxide composite*. In order to realize the full potential of additional charge storage, the thermal oxide has to be put under a higher electric field than is used in present day circuits. All the data on

breakdown improvement collected so far are obtained from fast ramp techniques. There may be some long term deteriorating effects with the thermal oxide under a higher stress field (up to 4 MV/cm) that cannot be observed by such fast testing techniques. Accelerated life test or long time life test has to be carried out to determine the full advantage of the composite structure.

VII. Conclusion

In conclusion, the advantages of the use of silicon-rich oxide and thermal oxide composite have been demonstrated. It can be used to increase storage capacitance in memory cells and the transconductance of FETs in present day circuits with a minimum perturbation in the processing. Furthermore, when the reliability and yield of thin gate dielectrics becomes a problem which occurs when scaling pushes device dimensions into the submicron regime, such a composite structure might be necessary.

ACKNOWLEDGEMENT

TABLE I

Wafer	Dielectric
A	10 nm thermal oxide
B	25 nm thermal oxide
C	35 nm thermal oxide
D	10 nm silicon-rich oxide/ 10 nm thermal oxide
E	20 nm silicon-rich oxide/ 10 nm thermal oxide
F	40 nm silicon-rich oxide/ 10 nm thermal oxide

REFERENCES

1. R.H. Dennard, "Field Effect Transistor Memory", U.S. Patent 3387286

2. T.C. May and M.H. Woods, "Alpha Particle Induced Soft Errors in Dynamic Memories", IEEE Trans. Electron Devices, **ED-26**, 2 (1979)

3. R.H. Dennard, F.H. Gaensslen, H.N. Yu, V.L. Rideout, E. Bassous and A.R. LeBlanc, "Design of Ion Implanted MOSFETs with Very Small Physical Dimensions", IEEE J. Solid St. Circuits, **SC-9**, 256 (1974)

4. M. Koyanagi, Y. Sakai, M. Ishihara, M. Tazunoki and N. Hashimoto, "A 5-V only 16-Kbit Stacked-Capacitor MOS RAM", IEEE Trans. Electron Devices, **ED-27**, 1596 (1980)

5. K. Ohta, K. Yamada, K. Shimizu and Y. Tarui, "Quadruply Self-Aligned Stacked High-Capacitance RAM Using Ta_2O_5 High Density VLSI Dynamic Memory", IEEE Trans. Electron Devices, **ED-29**, 368 (1982)

6. D. Dong, E.A. Irene and D.R. Young, "Preparation and Some Properties of Chemically Vapor-Deposited Si-Rich SiO_2 and Si_3N_4 Films", J. Electrochem. Soc., **125**, 819 (1978)

7. D.J. DiMaria and D.W. Dong, "High Current Injection into SiO_2 from Si-Rich SiO_2 Films and Experimental Applications", J. Appl. Phys., **51**, 2722 (1980)

8. D.J. DiMaria, R. Ghez and D.W. Dong, J. Appl. Phys., "Charge Trapping Studies in SiO_2 Using High Current Injection from Si-Rich SiO_2 Films", **51**, 4830 (1980)

9. D.J. DiMaria, K.M. DeMeyer, C.M. Serrano and D.W. Dong, "Electrically-Alterable Read-Only-Memory Using Si-Rich SiO_2 Injectors and a Floating Polycrystalline Silicon Storage Layer", J. Appl. Phys., **52**, 4825 (1981)

10. S. Yokoyama and S.K. Lai, unpublished results
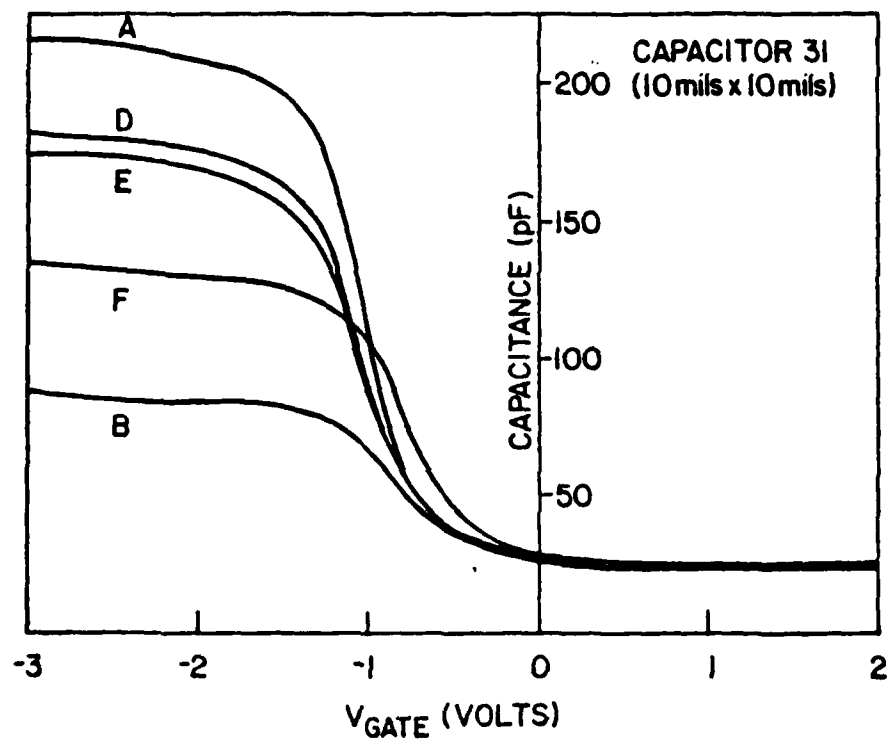
Figure 1. 1 MHz capacitance curves for square capacitors 10 mils on a side.
The thicknesses of the dual dielectric for the different wafers were
listed in Table I.
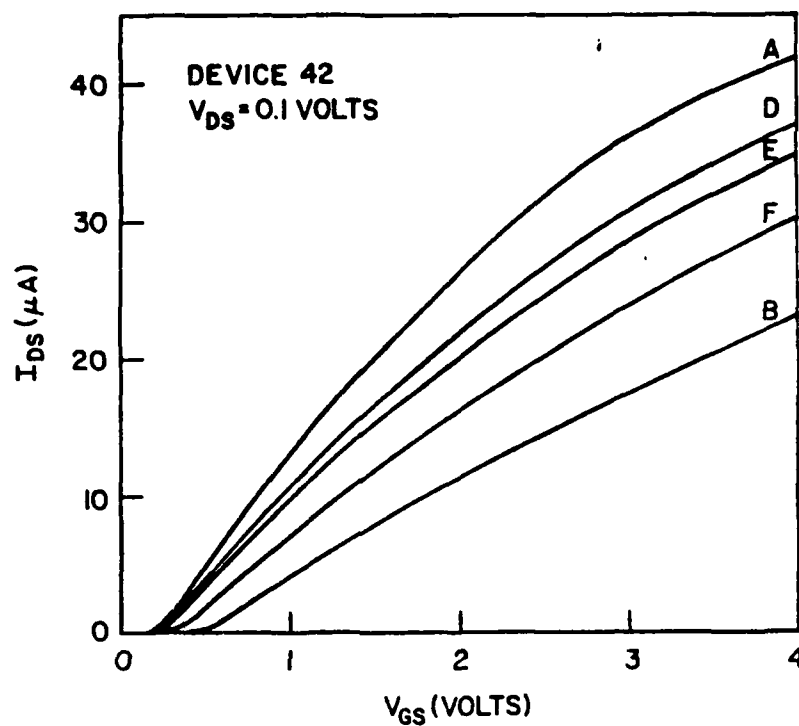
Figure 2. Source to drain currents as a function of gate voltage with source to drain voltages held at 100 mV. The source and substrate were grounded. Each curve is the average over eight devices located randomly over the wafer. The FETs have gate areas of approximately 4 mils by 4 mils and have channel width to length ratios of approximately one. Again, the thicknesses of the dual dielectric for the different wafers were listed in Table I.
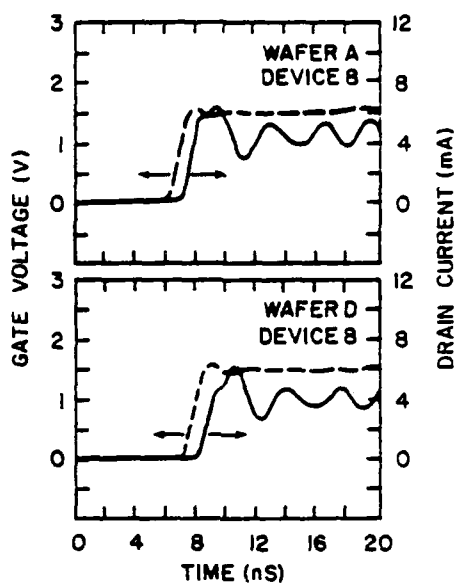
Figure 3. *Source to drain currents in response* to 0 to 1.5 volt steps applied to the gate. The source to drain voltages are held at 500 mV and the currents are sampled across a 10 ohm resistor. The substrate was grounded. In each frame, the dashed curve is for the gate pulse with sensitivity at 500 mV per division. The solid undulating curve is for the drain current with sensitivity at 20 mV per division. The undulations are due to reflectance because of improper termination. The horizontal scale is 2 nanosecond per division. The FETs have channel lengths of approximately 2 microns and large width to length ratios. The top trace was for device 8 on wafer A with 10 nm oxide. The bottom traces were for wafer D with 10 nm of silicon-rich oxide on top of 10 nm of thermal oxide.
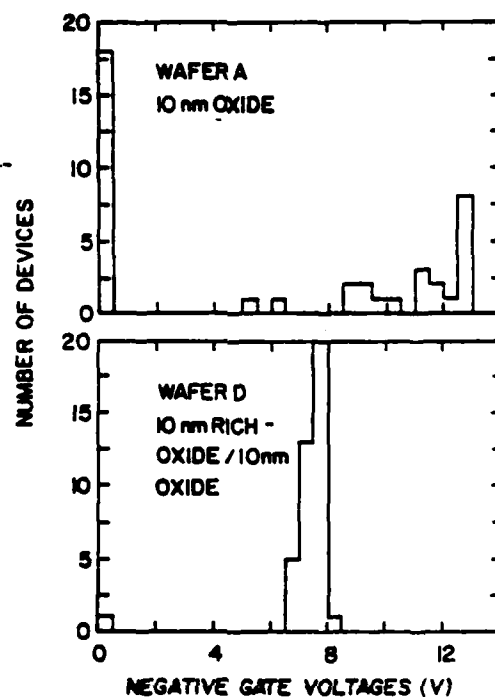
Figure 4. Breakdown histogram for 10 mil square capacitors. The breakdown current was set at approximately 0.16 Amp/cm^2 and the ramp rate was 0.5 MV/cm-sec.
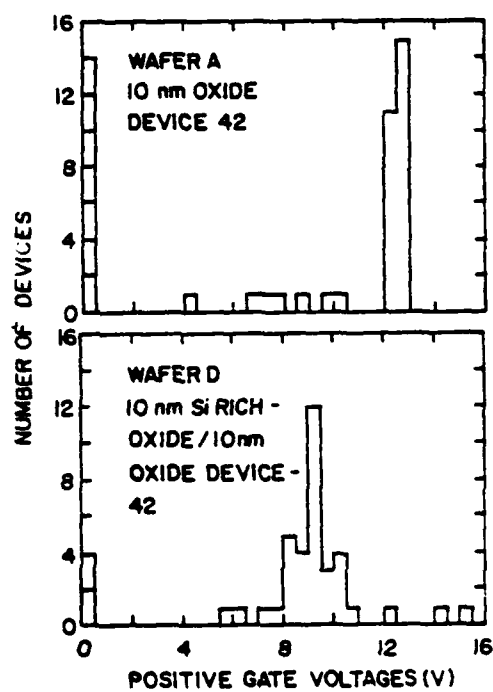
Figure 5. Breakdown histogram for FETs described in Figure 2. The source, drain and substrate are tied together. The breakdown current was set at approximately 1 Amp/cm^2 and the ramp rate was 0.5 MV/cm-sec.

RC 9970 (#44210) 4/29/83
Solid State Physics   37 pages

ENHANCED CONDUCTION AND MINIMIZED CHARGE TRAPPING IN ELECTRICALLY-ALTERABLE READ-ONLY-MEMORIES USING OFF-STOICHIOMETRIC SILICON DIOXIDE FILMS

D.J. DiMaria, D.W. Dong, and F.L. Pesavento
I.B.M. Thomas J. Watson Research Center
Yorktown Heights, New York 10598

and C. Lam
IBM GTD Burlington
Essex Junction, Vermont 05452

and S.D. Brorson
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Typed by L. D. Harris (ext.2985)

Abstract: An electrically-alterable read-only-memory using silicon dioxide and silicon-rich silicon dioxide layers capable of being cycled $\geq 10^7$ times by minimizing electron charge trapping in the SiO_2 layers of the device by incorporation of small amounts of silicon is discussed in detail. Charge transfer to and from a floating poly-crystalline silicon layer from a control gate electrode is accomplished by means of a modified dual-electron-injector-structure stack. This modified stack has the intervening silicon dioxide layer, which is sandwiched between silicon-rich silicon dioxide injectors, replaced by a slightly off-stoichiometric oxide containing between 1% and 6% excess atomic silicon above the normal 33% found in silicon dioxide. The operation of the electrically-alterable device structures in terms of write/erase voltages, cyclability, breakdown, and retention is related to current-voltage characteristics obtained from capacitors. A physical model based on direct tunneling between Si islands in the off-stoichiometric oxide layer is proposed to account for the observed increase in the moderate electric field conductance and decrease in charge trapping in these oxide layers incorporated into devices and capacitors. This model and the observed current-voltage characteristics are used to predict device operation for a variety of conditions.

I. Introduction

For years, many researchers in the field of solid state electronics and device physics have pursued the elusive non-volatile random-access-memory (NVRAM). This type of memory would be required to have information storage that could be changed very rapidly ($\lesssim$ 1 μsec), changed many times ($\gtrsim 10^{10}$ times), and hold the stored information for long periods of time ($\gtrsim$ 1 yr.) without being refreshed. A recent publication[1] gave some preliminary results indicating such a memory was evolving from a mixed phase oxide technology involving SiO_2 and Si-rich SiO_2 originally developed for a much slower memory ($\gtrsim$ 1 msec) which required less changes ($\lesssim 10^6$ times) called an electrically-alterable read-only-memory (EAROM)[2-5]. It is the purpose of this article to present detailed information on the physics and on the operation of these memory devices.

The original EAROM memory involved a field-effect-transistor (FET) configuration which had floating and control gate electrodes with a thermal SiO_2 layer from the single crystal Si substrate to the floating gate and a dual-electron-injector-structure (DEIS) in between the floating and control gates[2-5] as depicted in Fig. 1. This memory was called a DEIS EAROM, and the FET was turned on by the flow of electrons between source and drain contacts. Both the floating and control gates were formed from degenerately doped n-type polycrystalline silicon (poly-Si). The DEIS was formed from a layered stack of Si-rich SiO_2/SiO_2/Si-rich SiO_2 where the Si-rich SiO_2 layers had $\sim$ 13% excess atomic Si in them. The information storage was performed by putting electrons on the floating gate (called the "write" operation) or taking them off the floating gate (called the "erase" operation). The write and erase operations were accomplished at moderate average electric fields by field-enhanced electronic current injection at either the top or bottom Si-rich SiO_2/SiO_2 interfaces for negative and positive control gate voltages, respectively[2-5]. The localized field-enhancement was believed to be due to the two-phase nature of the Si-rich SiO_2 where Si islands in the SiO_2 matrix near the Si-rich SiO_2/SiO_2 interface injected electrons via a Fowler-Nordheim tunneling mechanism at lower average electric fields than a planar surface due to their finite curvature[6,7]. Since the Si islands were very small ($\lesssim$ 50 Å for Si-rich SiO_2 with $\sim$ 13% excess atomic Si) and densely packed compared to the smallest device areas ($\gtrsim 10^{-8} cm^2$), the current injection appeared to be uniform over the device area[6,7]. Since little voltage was dropped across the Si-rich SiO_2 layers during write or erase operations due to their large non-ohmic conductivity relative to the intervening SiO_2 layer at these electric fields[6-8], both the thickness of the Si-rich SiO_2 (from 100 Å to 1000 Å) and the exact percentage of excess atomic Si ($\gtrsim$ 13%) [9] were not critical for reproducible current-voltage characteristics. This makes the technology a very forgiving one. The information state of the

175

memory was determined during a "read" operation by sensing what control gate voltage was required to turn the FET on. These control gate voltages for reading were less than the voltages for writing or erasing[2-4,10]. If a net negative charge was on the floating gate electrode, a larger positive voltage than for the case of an uncharged floating gate was required to turn the FET "on" by forming an inversion layer of electrons at the Si-SiO$_2$ interface under this gate. This larger positive voltage was necessary to overcome the repulsive internal electric field created by the excess stored electrons on the floating gate electrode. For a net positively charged floating gate due to ionized donors, the opposite was true; that is, a smaller positive voltage was required due to the attractive internal electric field of the positive charges on the floating gate electrode.

The number of times the DEIS EAROM could be written and erased (or cycled) was determined by permanent electron capture on "deep" trapping states in the intervening SiO$_2$ layer of the DEIS stack[2,3,5]. The repulsive internal electric field created by the trapped electrons reduces the electric fields near the appropriate injecting Si-rich SiO$_2$/SiO$_2$ interface, thereby decreasing the injected current which depends strongly on this electric field. As the injected current is decreased, the electronic charge which is transferred between the floating and control gates is reduced, and the device no longer operates as originally specified after fabrication[3,5]. Since very little, if any, current passes between the floating gate and the Si substrate because the electric fields usually are not high enough for planar Fowler-Nordheim tunneling, charge trapping in the gate SiO$_2$ layer is not a problem[3]. These SiO$_2$ trapping sites are primarily due to H$_2$O related impurities in the bulk of the oxide with capture probabilities (that is, carriers captured to the total number injected) of between 10^{-3} to 10^{-6} and capture cross sections between 10^{-16} and 10^{-18}cm^{-2} [5,7,11-15]. Once captured on one of the sites, an electron can be detrapped optically with light of energy ≥ 4 eV[16] or thermally at temperatures $\geq 300°C$ in ambients containing hydrogen [3,12]. The number of water related impurities in the films can be reduced by extended high temperature annealing (1000°C in N$_2$) and careful processing[5,17]. However, traps with smaller capture cross sections ($\leq 10^{-19}$ cm^2) are still present in the SiO$_2$, regardless of the annealing or processing treatments[18]. Very little is known about these small cross section traps, and they may actually be created during extended current flow and/or with large electric fields across the SiO$_2$ layer. They do, however, ultimately limit the cyclability of DEIS EAROMs to $\leq 10^7$ cycles. Another approach which was tried to reduce the total number of SiO$_2$ electron traps and their effect involved making the SiO$_2$ layer very thin (40-60 Å) so that any remaining bulk-oxide trapped carriers could tunnel to the contacts. This approach did not prove to be very successful[19].

An approach which did prove to be successful in increasing cyclability dramatically to ~ 10^{10} cycles by minimizing the effects of permanent charge trapping involved adding a small amount of Si (1-6% excess atomic Si) to the intervening SiO_2 layer of the DEIS stack. Preliminary observations of increased cyclability out to at least $\geq 10^7$ cycles in EAROM devices using this modified DEIS (MDEIS) with an intervening off-stoichiometric (OS-SiO_2) layer have been reported recently[1]. The reason for the decrease in permanent electron trapping is believed to be due to a modified conduction mechanism in the OS-SiO_2 layer as depicted in Fig. 2. In the OS-SiO_2 films, most electrons injected from planar contacting electrodes or from Si-rich SiO_2 injectors with $\geq 13\%$ excess atomic Si move between the tiny Si islands in the OS-SiO_2 films rather than in the conduction band of the SiO_2 matrix. Evidence for this conduction mode, except at very large electric fields, using electroluminescence and p-channel FET carrier separation techniques have been recently reported[20]. Carriers tunneling directly between Si islands would be expected to be more difficult to capture than "free" electrons in the SiO_2 conduction band because of the reduction in available states (in particular, energetically shallow states) which mediate the capture process[11]. Also the Si islands could provide leakage paths for any electrons that are trapped into deep states. Under the appropriate electric field conditions, these trapped carriers could tunnel to nearby Si islands and than have the capability of leaving the OS-SiO_2 film by subsequent direct tunneling between Si islands. Although both positive and negative reversible space charges are observed to build-up near the contacting electrodes or Si-rich SiO_2 injectors under appropriate bias voltage conditions, hole currents as compared to electron currents through the OS-SiO_2 were determined to be at least a factor of 10^{-3} less[20]. Also compensation or annihilation of the permanent trapped electrons in the SiO_2 matrix of the OS-SiO_2 by positive charges on the Si islands or in the SiO_2 itself seems a doubtful explanation for the increase in device cyclability from previous experimental results[20] and from data which will be presented here.

In this article, emphasis will be given on relating the current-voltage (I-V) characteristics of OS-SiO_2 capacitors with or without Si-rich SiO_2 injector layers (see Fig. 3), to EAROM (or NVRAM) operating characteristics in terms of write/erase voltages and times, cyclability, retention of charges on the floating gate, and destructive breakdown of the oxide films. The dependence of these characteristics on Si content of the OS-SiO_2 films, voltage polarity, area, and temperature will be discussed in Section III. EAROM device modeling considerations using current-voltage characteristics as data input will be covered in Section IV.

II. Experimental

A) Sample Preparation

The off-stoichiometric oxide layers were deposited using atmospheric-pressure chemical-vapor-deposition (APCVD) techniques[8]. The excess silicon is introduced by adjusting the ratio (R_o) of the concentration of N_2O to SiH_4 in the gas phase. A 1% to 6% excess atomic silicon content in the oxide was obtained by setting the R_o value in the range of 50 to 30, as deduced from Rutherford backscattering (RBS). A stoichiometric SiO_2 film requires $R_o \geq 100$ while Si-rich SiO_2 injectors with an excess atomic Si content of 13% to 15% requires $R_o = 3$[6,8]. Film thickness and refractive index were deduced from ellipsometry measurements using techniques previously described[8]. The refractive index of these OS-SiO_2 films varied from 1.5 to 1.8 (1% to 6% excess atomic Si) while the low frequency dielectric constant varied from ~ 4 to 5 for the same range of excess Si. The samples used for ramp I-V and the point-by-point I-V measurements were metal-oxide-semiconductor (MOS) type capacitors involving layers of stoichiometric or off-stoichiometric oxides with or without Si-rich SiO_2 injectors. Structures with injectors sandwiching the SiO_2 or OS-SiO_2 layer are called DEIS or MDEIS, respectively. All oxides were deposited on <100> n-type or p-type silicon substrates of various resistivities (2, .2, .001 Ωcm) appropriate for the measurement. The different types of capacitor configurations are shown in Fig. 3. In most cases, the capacitor gate electrodes were formed from Al with a thickness of > 1000 Å and a circular area of .005 cm^2. For some capacitors, the gate electrode was formed from poly-Si which was degenerately doped n-type with $POCl_3$. For these capacitors the area of the poly-Si gate was varied from 4×10^{-7} cm^2 to 6.5×10^{-4} cm^2. All samples were annealed at 1000°C in an N_2 ambient for 30 min. before gate electrode deposition to reduce permanent electron charge trapping due to water related impurities [5,17]. In addition, a standard forming gas (90% N_2 /10% H_2) anneal at 400°C for 20 min. was performed on all samples after gate electrode depositions.

Figure 1 shows the basic type of EAROM device configuration used in this study. In some cases, devices were fabricated with equal-area floating and control gate electrodes or without the Si-rich SiO_2 injectors. For all EAROM device wafer runs, a set of capacitor monitors were fabricated where the SiO_2, OS-SiO_2, DEIS, or MDEIS was simultaneously deposited and processed (with Al gate electrodes) for I-V characterization. The EAROM devices were fabricated using a self-aligned, double polycrystalline silicon (poly-Si) gate process on 0.5 Ω-cm <100> p-type Si wafers. The details of the device fabrication process have been previously described[3,5]. Most of the device data presented here were taken on

3-port structures with a floating gate area of 2.5×10^{-6} cm^2 and a control gate area of 1.3×10^{-6} cm^2 (at the masking level) of n-degeneratively doped poly-Si[5]. These devices were square FETs with equal channel length and width (1.58×10^{-3} cm at the masking level), and they include XDEIS II 5-A,E,F,G,I-16 and DEIS II 2-C1-16. For MDT-NVR 1-E-7 devices, the floating gate and control gate areas were both equal to 8.4×10^{-6} cm^2. These devices were enclosed circular FETs with a channel length of 2.54×10^{-4} cm and a channel width of 1.62×10^{-2} cm^2 at the masking level. An extra masking step was involved in fabricating the devices with unequal control and floating gates. The gate oxides were thermally grown from the Si substrates at 1000°C in O_2 with 4.5% HCl to thickness of either 650 Å for the XDEIS II 5 and DEIS II 2 series of wafers or 100 Å for the MDT-NVR 1 series of wafers. The off-stoichiometric oxide layers in the MDEIS stacks (with or without injectors) were either 300 Å or 600 Å with varying amounts of excess Si. XDEIS II 5-E (with injectors) and F (no injectors) had 300 Å thick OS-SiO$_2$ with $R_o = 40$ (3-4% excess atomic Si.). XDEIS II 5-G, I and MDT-NVR 1-E all had 600 Å thick OS-SiO$_2$ with either $R_o = 40$ (3-4% excess atomic Si) for XDEIS II 5-G or $R_o = 30$ (5-6% excess atomic Si) for XDEIS II 5-I and MDT-NVR 1-E. The control DEIS EAROMS had stoichiometric intervening SiO$_2$ ($R_o = 200$) in the DEIS stack which were either 100 Å thick for DEIS II 2-C1 or 300 Å thick for XDEIS II 5-A. The Si-rich injectors with $\geq 13\%$ excess Si ($R_o = 3$) were either 200 Å in thickness for the XDEIS II 5-A,E,F,G,I and the MDT-NVR 1-E wafers or 150 Å in thickness for the DEIS II 2-C1 wafer. All OS-SiO$_2$, DEIS, or MDEIS layers were annealed in N_2 for 30 min. at 1000°C after deposition.

B)   Measurement Techniques

The ramp I-V measurements were performed by applying a voltage ramped at a constant rate ($|dV_g/dt| = 0.5$ V/sec) to the sample. The current through the capacitors was measured with a log-picoammeter (Keithley Model 26000). The temperature dependence of the ramp I-V characteristics was measured using a cryostat coupled with a liquid nitrogen transfer system (Air-products Heli-tran) for temperatures in the range of 80°K to room temperature (292°K). A resistance heated stage was used for temperatures above 22°C, up to 300°C. For current measurements where the voltage polarity could drive the Si substrate into deep depletion (in particular, p-Si substrates at large positive gate voltages), the samples were illuminated with white light. This insures rapid formation of an inversion layer at the substrate Si-oxide interface, negligible voltage drop across the Si substrate layer, and an adequate supply of carriers at this interface which can tunnel into the oxide. The white light generates minority carriers (electrons for p-Si) around the periphery of the opaque gate electrode which

flow laterally into the region under the electrode. However, this white light illumination had little effect on the conductivity of the off-stoichiometric oxides themselves over the electric field range of interest.

Point-by-point I-V measurements were performed by increasing the gate voltage magnitude in finite voltage steps and then measuring the current using a Keithley 616 digital electrometer after the capacitive current transient due to the voltage step had decayed. Constant current measurements were performed using a Keithley 417 fast picoammeter with current suppression and electronic circuits which continually readjusted the gate voltage to maintain the desired current level.

The voltage breakdown measurements were performed using a computer controlled automatic tester which ramped a capacitor from 0 V to a gate voltage necessary to draw the particle current desired, recorded this gate voltage, and then stepped the probe to the next capacitor. The voltage ramp rate magnitude was 5 V/sec and typically about 100 capacitors per wafer were tested. Each series of 100 capacitors per wafer were ramped to sequential current levels of 4×10^{-4} A/cm^2, 1×10^{-1} A/cm^2, and finally 4×10^{-4} A/cm^2 to test for breakdown and I-V reproducibility across the 1.25 inch diameter wafers. The doping of the Si substrate was picked so that the Si was in accumulation for all measurements; that is, n-type for positive gate voltages V_g^+ and p-type for negative gate voltages V_g^-. For all current measurements the Si substrate was held very close to ground potential (0 V).

The experimental arrangement for the EAROM FET measurements has been described in detail earlier[5]. The drain to source current was measured with a fast current amplifier (Keithley Model 127) while a ramp voltage was applied to the gate with the drain based at +0.1 V. Both gate voltage and drain to source current were displayed on a storage scope. The measurement was performed in fractions of a second to minimize read-disturb effects. The write/erase operation of the devices was carried out using square voltage pulses from two Hewlett-Packard pulse generators (Model 214A). The substrate was always held at ground potential (0 V). For retention measurements at elevated temperatures, a heated sample holder was used.

III Results and Discussion

In this section, experimental results from capacitor current-voltage characteristics using OS-SiO$_2$ films as a function of Si content of the oxide, area, temperature, voltage polarity, and the presence or/absence of Si-rich SiO$_2$ injector layers will be reviewed. Also, data on dielectric breakdown properties and I-V reproducibility on large area capacitors using OS-SiO$_2$

with or without Si-rich SiO_2 injectors will be presented and discussed. Then the EAROM device characteristics in terms of write/erase voltages and times, cycling, and retention using MDEIS stacks with OS-SiO_2 films will be discussed and related to the I-V characteristics.

A. Capacitors

1. Dark Currents

The current I as a function of voltage characteristics for off-stoichiometric oxides incorporated into DEIS stacks with Si-rich SiO_2 injectors (see Fig 3b) are illustrated in Fig. 4 for APCVD oxides with [N_2O] to [SiH_4] ratios of R_o=200, 50, 40 and 30. This ratio R_o=200 corresponds to a stoichiometric oxide while R_o=30, 40, and 50 corresponds to approximately 5-6%, 3-4%, 1-2% excess atomic silicon in the oxide, respectively, as deduced from Rutherford backscattering measurements. The ramp I-V plots shown in Figs. 4-7 were generated by applying a constant rate voltage ramp ($|dV_g/dt| = 0.5$ V/sec) to the samples up to a measured current on the order of 10^{-5} to 10^{-4} A; the voltage ramp was then reversed at the same rate. A displacement current ($I_D = C \cdot dV_g/dt$, where C is the capacitance associated with the MOS structures) shows as an approximately constant current component at low applied electric fields. Current as a function of voltage characteristics for off-stoichiometric oxides with increasing Si content incorporated into MOS capacitors without Si-rich SiO_2 injectors (see Fig 3a) showed a behavior similar to the data of Fig. 4, but with all curves moved to large gate voltages. This is shown in Fig. 5. This hysterisis observed in the I-V characteristics which is dependent on voltage ramp rate is associated with permanent charge trapping in the oxide[7]. Since permanent charge trapping occurs in the <u>bulk</u> of the SiO_2 [7,11,13,15], reduction in I-V hysterisis similar to that in Figs. 4 and 5 is observed for both negative and positive gate voltages (V_g^- and V_g^+). Decreasing the amount of the SiO_2 in the two phase matrix of the off-stoichiometric oxide by increasing the excess Si should decrease the amount of electrons trapped in the SiO_2 regions and therefore the hysterisis. Typically thinner stoichiometric SiO_2 layers in MOS capacitors give less hysterisis in ramp I-V characteristics. DEIS or MOS configurations (see Fig. 3) with little or no hysterisis (for example, the R_o=30 DEIS in Fig. 4) had I-V characteristics which were independent of the ramp rate from .005 V/sec to 5 V/sec. This hysterisis did not recover after periods of days with the gate electrode floating or grounded on off-stoichiometric oxides in DEIS stacks (Fig. 4) or MOS capacitors with no injectors (Fig. 5). From these I-V, data two important effects are observed as the excess silicon in the oxides is increased (R_o is decreased):

i) the permanent trapped charge in the oxide decreases, and ii) lower electric fields are required to induce charge transport across the oxide layers.

Figure 6 compares MDEIS stacks with off-stoichiometric oxides in MOS capacitor configurations (See Fig. 3b) with MOS capacitors without the Si-rich SiO_2 injector layers of the MDEIS (See Fig 3a). As seen in this figure, the MDEIS current-voltage characteristic is moved to lower gate voltage magnitudes as compared to the off-stoichiometric oxide alone. This implies the MDEIS structure with the intervening off-stoichiometric oxide is still somewhat sensitive to the localized electric field enhancement near the Si-rich SiO_2 injector interface with the intervening oxide layer, and that these interfaces must influence and limit the currents measured in the external circuit somewhat.

Current-average electric field characteristics of capacitor structures with or without Si-rich SiO_2 injectors were observed to show a weak dependence on:

1.) contact electrode material, including Si, Al, or Au ;

2.) gate voltage polarity, especially when compared to stoichiometric oxide or DEIS capacitors ;

3.) time in the range from 500 μsec to 20 sec ;

4.) OS-SiO_2 thickness in the range from 300-1200 Å ;

5.) gate electrode area (or perimeter) in the range from 4×10^{-7} cm^2 to 6.5×10^{-4} cm^2 as shown in Fig. 7;

6.) and temperature in the range from 77°K (liquid nitrogen) to 300°C, except at very low average electric fields ($\lesssim 1.5$ MV/cm) as shown in Fig. 8.

Observations 1 and 2 imply that a reversible space charge layer of finite thickness must build up in the oxide in front of the injecting electrode on near an injector-oxide interface. Observation 3 implies this space charge layer builds up rapidly, while observation 4 implies that it must have a thickness < 300 Å in extent. Observation 5 implies that the injection into and conduction through the off-stoichiometric oxide is uniform in a macroscopic sense for the scale of the device and capacitor areas used here (larger than microns by microns). The lack of a strong temperature dependence (observation 6) is consistent with quantum mechanical

182

tunneling controlling and limiting the currents measured in the external circuits as will be discussed next.

The I-V data in Figs. 4-7 have a very strong dependence on the average electric field E. The average electric field defined as $E = (V_g - \Phi_{ms} - \Psi_s)/\ell_o$ where V_g is the voltage applied to the gate electrode, Φ_{ms} is the difference in work function between the gate material and the Si substrate (by convention, expressed in volts), Ψ_s is the surface potential for the substrate Si-oxide interface, and ℓ_o is the oxide thickness.

The functional form of the current density J $(J=I/A)$ where A is the injection gate area) can be expressed as

$$J = a\, f(E)\, \exp\,[-b/E] \qquad (1)$$

where a and b are constants and $f(E) = E^n$ (n = 1 or 2)[20-22]. This type of expression for J with n=1 is that expected for tunneling directly between Si islands in the OS-SiO$_2$ films through a trapezodial energy barrier films at low to moderate electric fields[20,22]. With n=2, Eq. 1 is equivalent to the Fowler-Nordheim expression for tunneling through a triangular energy barrier from near the bottom of the conduction band of a Si island to the bottom of the conduction band of the SiO$_2$ matrix at very high electric fields[20,21]. Good least squares fits to the I-V data for OS-SiO$_2$ films with or without Si-rich SiO$_2$ injectors, such as the data in Figs. 4-7, was obtained using either value of n. In Section IV, Eq. 1 with n=2 will be used to deduce values of a and b from the experimental ramp I-V data in order to model and predict the operation of the EAROMs which will be discussed in Section III-B.

2. Breakdown

Statistical studies of ramped current-voltage characteristics on 100 capacitors of MDEIS and control DEIS stacks with intervening stoichiometric SiO$_2$ deposited on 2 Ωcm, <100> orientation, p-Si substrates were performed. The MDEIS and DEIS stacks used in the large area (.005 cm^2) circular capacitors with Al gate electrodes were deposited under the same conditions as some of the corresponding small area FET devices described in later sections. All 100 capacitors per wafer were ramped at -5 V/sec (applied to the gate) to sequential current levels of 4×10^{-4} A/cm^2, 1×10^{-1} A/cm^2, and 4×10^{-4} A/cm^2 and the voltages at which these current levels were reached were recorded. For 100 capacitors, the subsequent histograms showed $\leq$

183

10% of the capacitors were shorted initially (mostly near the edges of the wafer) and $\lesssim$ 5% new breakdowns were added with the high current cycling sequence. The remaining $\approx$90% of the devices just demonstrated the current-voltage reproducibility across the 1.25 inch diameter wafers. These results are consistent with the breakdown improvement (in particular, the minimization of the low field destructive breakdowns) reported before for DEIS stacks with stoichiometric SiO_2 layers[3].

Figures 9 and 10 show some of the histograms with .25 MV/cm bins for this testing sequence on MDEIS capacitors fabricated with the XDEIS II 5 series of EAROM wafers with 200 Å thick Si-rich SiO_2 injectors and a 600 Å thick intervening OS-SiO_2 layer with either 3-4% excess atomic Si (Fig. 9) or 5-6% excess atomic Si (Fig. 10). Clearly the I-V rechecks in these figures (frame C) are very similar to the initial I-V values (frame A) after ramping to high current and electric field (frame B). Some shift of the histogram recheck run to larger electric field magnitudes was observed on the control DEIS capacitors due to permanent charge trapping in the stoichiometric SiO_2 layer during the total voltage ramping sequence. Similar break-down improvement was observed on capacitors with just the OS-SiO_2 layer and no Si-rich SiO_2 injectors, and on capacitors with OS-SiO_2 layers deposited on n-type Si substrates which were ramped under positive gate voltage conditions to the same current levels in the same sequence. This breakdown improvement was seen on capacitors with OS-SiO_2 layers down to thicknesses of 300 Å. However, capacitors with $\leq$ 150 Å thick OS-SiO_2 layers with or without Si-rich SiO_2 injectors were usually more likely to suffer destructive breakdown of the dielectric at lower average electric fields. As discussed in previous publications, the low field breakdown improvement is believed to be due to electric field screening caused by the reversible space charge that builds-up on Si islands in the insulator layers near the contacts in either the Si-rich SiO_2 injectors[3], the OS-SiO_2, or possibly both. This space charge screens the effects of any localized high field spots at the contact/oxide interface due to asperities, particulate, or defects which would locally cause a high current density injection. Contact/oxide interface irregularities are believed to be a leading cause of destructive breakdown of the insulator at low average electric fields[23].

The ultimate breakdown strength of insulators such as SiO_2 is believed to be related to the electric field strength[24]. For DEIS capacitors, catastrophic breakdown usually occurs when electric fields exceed ~9 MV/cm[3,7]. Breakdown properties of DEIS, MDEIS, and OS-SiO_2 capacitors were tested using constant current conditions. In these experiments, the capacitors were held under constant current conditions

$(6 \times 10^{-3}$ to 1×10^{-1} A/cm^2) using circuits which automatically adjusted the gate voltage to compensate for changes in the electric field near the injecting interface due to electron trapping in the SiO$_2$ regions. Current and gate voltage were monitored as a function of time until destructive breakdown occurred. For DEIS structures, breakdown usually occurred after bulk SiO$_2$ electron trapping forced electric fields to be > 9 MV/cm. For the MDEIS or OS-SiO$_2$ capacitors (such as those in Figs. 9 and 10) which operate at lower average electric fields and have much less charge trapping than the DEIS capacitors, destructive breakdown was usually not observed to occur under similar constant current conditions and stressing times. Typically DEIS structures broke down destructively after passing $\leq$ 10 coul/cm^2 of electronic charge while MDEIS or OS-SiO$_2$ structures showed no signs of breakdown after passing $\geq$ 100 coul/cm^2 of electronic charge.

B. EAROM Devices

1. Write/Erase Operation

The use of off-stoichiometric oxides (1-6% excess atomic Si) instead of stoichiometric SiO$_2$ in the intervening layer of a DEIS EAROM structure results in extended cycling characteristics for these devices, as will be demonstrated in this section. A schematic representation of a DEIS/MDEIS EAROM is shown in Fig. 1. Electronic charge is transferred back and forth from the control gate of area A_1 to the floating gate of area A_2 by means of enhanced injection and conduction of the MDEIS stack with its intervening off-stoichiometric oxide layer. The injector regions of the MDEIS or DEIS structures contain $\geq$ 13% excess atomic Si compared to the 1% to 6% excess atomic Si used in the intervening oxide layer. Very little of the applied voltage is dropped across the injector regions during write or erase operations. Writing or erasing (that is, putting negative charge on or taking it off the floating gate) is achieved by applying a negative or positive voltage to the control gate electrode, respectively[3]. The threshold voltage V_T (that is, the gate voltage at which the channel of the device turns on[25]) of the DEIS or MDEIS EAROMs is used as a measure of the charge state (written or erase) of the floating gate electrode. More specifically, the threshold voltage shift

$$\Delta V_T = V_T(t) - V_{T_i} \tag{2}$$

where t is the write/erase pulse time duration and V_{T_i} is the threshold voltage for a device with an uncharged floating gate, or the change in the threshold voltage shift

$$\Delta[\Delta V_T] = \Delta V_T(t) - \Delta V_T(0) = V_T(t) - V_T(0) \tag{3}$$

are the quantities used as indicators of the charge on the floating gate. For write operations with negative gate voltages, both ΔV_T and $\Delta[\Delta V_T]$ are positive quantities indicating that electrons are stored on the floating gate during the write pulse. For the erase operation with positive gate voltages, both ΔV_T and $\Delta[\Delta V_T]$ are negative quantities indicating that a net positive charge (ionized donors) was left on the floating gate after the erase pulse. More specifically, the charge transferred $\Delta Q = Q(t) - Q(0)$ can be obtained from the relationship[4]

$$\Delta Q = -C_1^* \Delta[\Delta V_T] = -C_1^* [\Delta V_T(t) - \Delta V_T(0)] \tag{4}$$

where $C_1^* = A_1(\ell_{o_1}/\varepsilon_{o_1} + \ell_n/\varepsilon_n)^{-1}$ is the <u>low</u> <u>field</u> capacitance of the DEIS or MDEIS stack which includes the series contribution of both Si-rich SiO_2 injectors which have a low-field, low-frequency permittivity of $\varepsilon_n \approx 7.5 \times$ (permittivity of free space)[3,4,7,10] and a total sum of thicknesses $\ell_n = \ell_{n_1} + \ell_{n_2}$. The intervening oxide thickness of the DEIS or MEIS is defined by ℓ_{o_1} with a low-frequency permittivity of $\varepsilon_{o_1} \approx (3.9\text{-}6) \times$ (permittivity of free space) depending on the Si content from $R_o = 50$ films with 1-2% excess atomic Si to $R_o = 30$ films with 5-6% excess atomic Si, respectively (see section II-A). C_1^* is used because the threshold voltage of the devices is determined at applied control gate voltages whose magnitude is usually small compared to the write/erase voltages. During write or erase $C_1^* \rightarrow C_1 = \varepsilon_{o_1} A_1 / \ell_{o_1}$ due to the increased conductivity of the Si-rich SiO_2 injectors with respect to the intervening oxide layer which effectively force ε_n to a large value[3,4,7,10]. However, if the Si-rich injectors are thin (<200 Å) with high Si content ($>50\%$ atomic Si) or if the floating gate is sufficiently overcharged, $C_1^* \rightarrow C_1$ even during the read operation. Of course for $OS\text{-}SiO_2$ EAROMS without Si-rich SiO_2 injectors, $C_1^* = C_1$. Another important quantity which will be used in latter sections is the average electric field across the intervening oxide layer of the DEIS or MDEIS during the write or erase operation, which is defined by[4]

$$E_{o_1}(t) = \frac{\chi(V_g - \Phi_{ms} - \Psi_s) + C_1^* \Delta V_T(t)}{\ell_{o_1} C_T} \tag{5}$$

where $\chi = C_2 = \varepsilon_{o_2} A_2 / \ell_{o_2}$ and $C_T = C_1 + C_2$ for the 3 port structures discussed here. The gate SiO_2 thickness from the floating gate to the Si substrate is defined by ℓ_{o_2} with a low-frequency permittivity of $\varepsilon_{o_2} = 3.9 \times$ (permittivity of free space).

Figures 11-15 show data for the change in the threshold voltage shift as a function of the write or erase voltage applied to the control gate for times varying from 5 msec to 500 nsec where the initial charge state of the floating gate was set by the condition $V_T(0) = 1.7$ V. The conclusions which can be obtained from these data are the following:

1.) OS-SiO$_2$ MDEIS EAROMs with increasing Si content leads to lower write/erase voltage magnitudes (see Fig. 11).

2.) OS-SiO$_2$ EAROMs without Si-rich SiO$_2$ injectors require larger write/erase voltage magnitudes than comparable MDEIS EAROMs with injectors and have greater asymmetry between the write and erase voltage conditions (see Fig. 12). This was also seen for any write/erase pulse time from 5 msec to 500 nsec.

3.) OS-SiO$_2$ MDEIS EAROMs are easier to write or erase than comparable DEIS EAROMs with stoichiometric intervening SiO$_2$ layers for long pulse times ($\gtrsim$ 500 μsec), but are more difficult to charge for fast pulsing (compare Figs. 13 and 14).

4.) Thinner OS-SiO$_2$ EAROMs are easier to write than thicker ones (compare Figs. 11, 14, and 15).

These observations on the EAROM devices are predictable from the ramp I-V measurements discussed in section III-A-1. For example, conclusion 1 is consistent with Figs. 4 and 5. Conclusion 2 is consistent with Fig. 6, even after corrections for changes in the low field capacitance of the MDEIS stack C_1^* due to the absence of the injector layers (see Eq. 4). Conclusion 3 occurs due to the OS-SiO$_2$ oxides becoming very conductive at high electric fields. At very high fields, $J \approx af(E)$ because the tunneling barrier becomes almost transparent; that is, the exponential term in the current density relationship approaches ~ 1 (See Eq. 1). This will be discussed in more detail in the Section IV (device modeling), and it imposes certain limitations on write/erase speed. The roll-off and flattening of the $|\Delta[\Delta V_T]|$ vs. $|V_g|$ data in Fig. 14 is due to the over-charging of the floating gate. For this case after the write or erase pulse has ended and the control gate is returned to ground, the internal electric field between the floating and the control gates is large enough to either conduct electrons off the floating gate to the control gate or vice-versa, respectively.

187

Figures 16 and 17 demonstrate space charge effects in the DEIS and MDEIS stacks of EAROM devices, respectively. The space charge effects in the DEIS EAROMs have been reported previously[26]. These effects are believed to be associated with the reversible trapped space charge which builds up on the Si islands in the Si-rich SiO_2 injector layers of the DEIS stack. This reversible space charge due to trapped electrons degrades the device operation by building up near the injecting contact, screening the electric field, and thereby lowering the injection efficiency from the contact. This phenomenon is more pronounced if the Si-rich SiO_2 layer is thick and/or has a small percentage of excess Si so that these trapped electrons can not move as easily through the Si-rich SiO_2 layer to the Si-rich-SiO_2/SiO_2 interface and then be injected into the SiO_2 layer[3,26]. In Fig. 16, a longer delay between each pulse of a string of write or erase pulses gives a larger threshold voltage shift magnitude since more electrons are put onto or removed from the floating poly-Si layer. This is because the trapped electrons on the Si islands which block contact injection have had sufficient time to move back to the appropriate contact possibly having this motion influenced by their own internal electric field. Fig. 16 also shows that a large number of pulses of small duration are more efficient than one large pulse of the same total time duration. Again this is due to the trapped electronic charge build-up on the Si islands limiting contact injection particularly when there are few delays, delays of small duration, or no delays between each pulse of the write or erase pulse train. Data similar to Fig. 16 are also observed for write/erase voltage conditions where smaller amounts of charge are transferred, indicated by smaller $|\Delta[\Delta V_T]|$ values, on XDEIS II 5-A-16 devices (the DEIS has a 300 Å intervening SiO_2 layer).

The data in Fig. 17 for an MDEIS EAROM shows a trend in the opposite direction as compared to that for the DEIS EAROM shown in Fig. 16. Here a single large 5 msec pulse is more efficient in charging or discharging the floating gate than a pulse train of shorter duration pulses adding to a total pulsing time of 5 msec. Also less charge is transferred for increasing delay time between the pulses of a write or erase pulse train. Data similar to that in Fig. 17 were observed for different amounts of transferred charge on this device. Other devices with different MDEIS stacks or just OS-SiO_2 (no injectors) between the floating and control gates, regardless of composition (1-6% excess atomic Si) or OS-SiO_2 thickness (300-600 Å), showed data similar to Fig. 17 for various amounts of transferred charge. Increasing the excess Si content in the OS-SiO_2 layer of the devices, decreased the magnitude of the differences observed for the various pulse trains as compared to a single 5 msec pulse.

Devices with injectors (MDEIS EAROMS) had smaller differences than devices without injectors for the various pulsing sequences. The MDEIS is believed to operate differently than the DEIS because of subtle changes in the spatial extent of the space charge on Si islands in the intervening OS-SiO$_2$ oxide (1-6% excess atomic Si) near the Si-rich SiO$_2$ injecting layers ($\geq$ 13% excess atomic Si). For the MDEIS, the space charge is not confined just to the injector, but it moves into the OS-SiO$_2$ layer in time. The deeper spatially this space charge moves into the OS-SiO$_2$, the higher the local electric fields in the SiO$_2$ matrix portion of the remaining bulk OS-SiO$_2$ where there is probably little or no space charge on the Si islands. Higher fields give increased conductivity, and therefore more charge is transferred to the floating gate. In more conductive OS-SiO$_2$ films, the effect demonstrated by the data in Fig. 17 would be less pronounced because a steady state condition for the space charge regions would be reached quicker and the extent of penetration of the space charge would be less in the OS-SiO$_2$ layer.

2. Cycling

One of the main conclusions of this work and the advantage of MDEIS over DEIS EAROMs is demonstrated in Figs. 18 and 19. In these figures, the MDEIS EAROMs with OS-SiO$_2$ can be cycled orders of magnitude more times than comparable DEIS EAROMs with SiO$_2$ without degradation of write/erase operation. For these cycling data, devices were written from an erased state or erased from a written state. These figures show the threshold voltage for the written V_{T_W} and erased V_{T_E} states as a function of the number of write/erase cycles for similar EAROM devices involving stoichiometric SiO$_2$ ($R_0 = 200$) and off-stoichiometric oxides as intervening layers in DEIS or MDEIS stacks, respectively. The threshold voltage difference between the written and the erased states $\Delta[\Delta V_T] = V_{T_W} - V_{T_E}$ (called the threshold voltage window) starts to collapse in $\sim 10^3$ cycles on the devices with stoichiometric SiO$_2$ while a similar collapse starts to occur after $\sim 10^7$ cycles on devices with off-stoichiometric oxides. The improvement in the cycling characteristics of these devices results from the reduction in the effect of permanent electron trapping in the off-stoichiometric oxide layer. As some of the transferred electrons are permanently trapped, they create a repulsive internal electric field which lowers the electric field near the injecting interface which in turn controls the current flow to and from the floating gate. As the interface field decreases, fewer electrons for the same write or erase voltage condition can be transferred and the threshold voltage difference between the written and erased states (which is a measure of the transferred charge)

189

Figures 11-15 show data for the change in the threshold voltage shift as a function of the write or erase voltage applied to the control gate for times varying from 5 msec to 500 nsec where the initial charge state of the floating gate was set by the condition $V_T(0) = 1.7$ V. The conclusions which can be obtained from these data are the following:

1.)   OS-SiO$_2$ MDEIS EAROMs with increasing Si content leads to lower write/erase voltage magnitudes (see Fig. 11).

2.)   OS-SiO$_2$ EAROMs without Si-rich SiO$_2$ injectors require larger write/erase voltage magnitudes than comparable MDEIS EAROMs with injectors and have greater asymmetry between the write and erase voltage conditions (see Fig. 12). This was also seen for any write/erase pulse time from 5 msec to 500 nsec.

3.)   OS-SiO$_2$ MDEIS EAROMs are easier to write or erase than comparable DEIS EAROMs with stoichiometric intervening SiO$_2$ layers for long pulse times ($\gtrsim$ 500 μsec), but are more difficult to charge for fast pulsing (compare Figs. 13 and 14).

4.)   Thinner OS-SiO$_2$ EAROMs are easier to write than thicker ones (compare Figs. 11, 14, and 15).

These observations on the EAROM devices are predictable from the ramp I-V measurements discussed in section III-A-1. For example, conclusion 1 is consistent with Figs. 4 and 5. Conclusion 2 is consistent with Fig. 6, even after corrections for changes in the low field capacitance of the MDEIS stack C_1^* due to the absence of the injector layers (see Eq. 4). Conclusion 3 occurs due to the OS-SiO$_2$ oxides becoming very conductive at high electric fields. At very high fields, $J \approx af(E)$ because the tunneling barrier becomes almost transparent; that is, the exponential term in the current density relationship approaches ~ 1 (See Eq. 1). This will be discussed in more detail in the Section IV (device modeling), and it imposes certain limitations on write/erase speed. The roll-off and flattening of the $|\Delta[\Delta V_T]|$ vs. $|V_g|$ data in Fig. 14 is due to the over-charging of the floating gate. For this case after the write or erase pulse has ended and the control gate is returned to ground, the internal electric field between the floating and the control gates is large enough to either conduct electrons off the floating gate to the control gate or vice-versa, respectively.

187

Figures 16 and 17 demonstrate space charge effects in the DEIS and MDEIS stacks of EAROM devices, respectively. The space charge effects in the DEIS EAROMs have been reported previously[26]. These effects are believed to be associated with the reversible trapped space charge which builds up on the Si islands in the Si-rich SiO_2 injector layers of the DEIS stack. This reversible space charge due to trapped electrons degrades the device operation by building up near the injecting contact, screening the electric field, and thereby lowering the injection efficiency <u>from the contact.</u> This phenomenon is more pronounced if the Si-rich SiO_2 layer is thick and/or has a small percentage of excess Si so that these trapped electrons can not move as easily through the Si-rich SiO_2 layer to the Si-rich-SiO_2/SiO_2 interface and then be injected into the SiO_2 layer[3,26]. In Fig. 16, a longer delay between each pulse of a string of write or erase pulses gives a larger threshold voltage shift magnitude since more electrons are put onto or removed from the floating poly-Si layer. This is because the trapped electrons on the Si islands which block contact injection have had sufficient time to move back to the appropriate contact possibly having this motion influenced by their own internal electric field. Fig. 16 also shows that a large number of pulses of small duration are more efficient than one large pulse of the same total time duration. Again this is due to the trapped electronic charge build-up on the Si islands limiting contact injection particularly when there are few delays, delays of small duration, or no delays between each pulse of the write or erase pulse train. Data similar to Fig. 16 are also observed for write/erase voltage conditions where smaller amounts of charge are transferred, indicated by smaller $|\Delta[\Delta V_T]|$ values, on XDEIS II 5-A-16 devices (the DEIS has a 300 Å intervening SiO_2 layer).

The data in Fig. 17 for an MDEIS EAROM shows a trend in the opposite direction as compared to that for the DEIS EAROM shown in Fig. 16. Here a single large 5 msec pulse is more efficient in charging or discharging the floating gate than a pulse train of shorter duration pulses adding to a total pulsing time of 5 msec. Also less charge is transferred for increasing delay time between the pulses of a write or erase pulse train. Data similar to that in Fig. 17 were observed for different amounts of transferred charge on this device. Other devices with different MDEIS stacks or just OS-SiO_2 (no injectors) between the floating and control gates, regardless of composition (1-6% excess atomic Si) or OS-SiO_2 thickness (300-600 Å), showed data similar to Fig. 17 for various amounts of transferred charge. Increasing the excess Si content in the OS-SiO_2 layer of the devices, decreased the magnitude of the differences observed for the various pulse trains as compared to a single 5 msec pulse.

Devices with injectors (MDEIS EAROMS) had smaller differences than devices without injectors for the various pulsing sequences. The MDEIS is believed to operate differently than the DEIS because of subtle changes in the spatial extent of the space charge on Si islands in the intervening OS-SiO_2 oxide (1-6% excess atomic Si) near the Si-rich SiO_2 injecting layers ($\geq$ 13% excess atomic Si). For the MDEIS, the space charge is not confined just to the injector, but it moves into the OS-SiO_2 layer in time. The deeper spatially this space charge moves into the OS-SiO_2, the higher the local electric fields in the SiO_2 matrix portion of the remaining bulk OS-SiO_2 where there is probably little or no space charge on the Si islands. Higher fields give increased conductivity, and therefore more charge is transferred to the floating gate. In more conductive OS-SiO_2 films, the effect demonstrated by the data in Fig. 17 would be less pronounced because a steady state condition for the space charge regions would be reached quicker and the extent of penetration of the space charge would be less in the OS-SiO_2 layer.

2. Cycling

One of the main conclusions of this work and the advantage of MDEIS over DEIS EAROMs is demonstrated in Figs. 18 and 19. In these figures, the MDEIS EAROMs with OS-SiO_2 can be cycled orders of magnitude more times than comparable DEIS EAROMs with SiO_2 without degradation of write/erase operation. For these cycling data, devices were written from an erased state or erased from a written state. These figures show the threshold voltage for the written V_{T_W} and erased V_{T_E} states as a function of the number of write/erase cycles for similar EAROM devices involving stoichiometric SiO_2 (R_o=200) and off-stoichiometric oxides as intervening layers in DEIS or MDEIS stacks, respectively. The threshold voltage difference between the written and the erased states $\Delta[\Delta V_T] = V_{T_W} - V_{T_E}$ (called the threshold voltage window) starts to collapse in ~ 10^3 cycles on the devices with stoichiometric SiO_2 while a similar collapse starts to occur after ~ 10^7 cycles on devices with off-stoichiometric oxides. The improvement in the cycling characteristics of these devices results from the reduction in the effect of permanent electron trapping in the off-stoichiometric oxide layer. As some of the transferred electrons are permanently trapped, they create a repulsive internal electric field which lowers the electric field near the injecting interface which in turn controls the current flow to and from the floating gate. As the interface field decreases, fewer electrons for the same write or erase voltage condition can be transferred and the threshold voltage difference between the written and erased states (which is a measure of the transferred charge)

189

decreases. However, although an improvement of $\gtrsim 4$ orders of magnitude in the number of cycles is achieved, the threshold voltage window still collapses (see Fig. 19). This implies that permanent charge trapping in the oxide is still present. It is possible that a small percent of the electrons are still transported through the intervening oxide layer via the bottom of the conduction band of the SiO_2 phase giving rise to electron capture in deep traps ($\gtrsim 4$ eV from the bottom of the SiO_2 conduction band) and/or that not all electrons trapped in the SiO_2 phase of the OS-SiO_2 can tunnel to nearby Si islands as was discussed in Section I.

The permanent charge trapping efficiency is further reduced as the excess silicon content in the oxide is increased or the OS-SiO_2 is made thinner (see Fig. 18). This is consistent with reduction in ramp I-V hysterisis due to the reduction of permanent electron trapping in OS-SiO_2 films compared to SiO_2 as discussed in Section III-A-1 (see Figs. 4 and 5). Figure 19 also shows that the cycling collapse is not as abrupt with the MDEIS as compared to the DEIS EAROM and that $> 10^{10}$ cycles can be achieved.

The threshold voltage window collapse in the DEIS EAROM with the stoichiometric SiO_2 intervening layer of Figs. 18 and 19 has been accelerated by design. This is achieved by using thicker intervening SiO_2 layers (300 Å) than used in devices reported in previous publications[2-5] to get more bulk SiO_2 electron trapping. This is apparent if the rate of threshold voltage window collapse as a function of oxide thickness on the DEIS EAROMs is compared for SiO_2 thickness in the range from 100 Å to 600 Å. However the best DEIS cyclability with thin SiO_2 would show complete collapse of the threshold voltage window by $\leq 10^7$ cycles[5].

In Fig. 18, the write/erase pulse duration for the cycling experiments was 500 μsec while in Fig. 19 it was 500 nsec. Accordingly, the write/erase voltages increased in a manner consistent with previous discussions (see Figs. 13-15). To get lower control gate voltages with faster write/erase speeds as shown in Fig. 20, the gate SiO_2 layer from the Si substrate to the floating gate was made thinner (100 Å). This gate voltage magnitude spread from 13 V to 34 V for write/erase times from 5 msec to 5 μsec is quite reasonable. These particular devices (MDT-NVR 1 wafers) have equal floating and control gate areas. Further reduction in the write/erase gate voltage magnitude could be achieved with unequal gate area devices where the control gate is smaller than the floating gate, such as with the XDEIS mask design. However, there is a limit to how low the write/erase voltages can be made for single device EAROM

cells like those discussed here before the voltage associated with the read operation disturbs the charge stored on the floating gate during the write or erase operation. Two-device cells which have an additional FET in series with the DEIS EAROM will minimize these "read disturb" effects and allow even lower voltages than those discussed here. By addressing this addition FET with the DEIS EAROM gate grounded, the cell can be read by sensing current flow (erased state) or no current flow (written state) through the two devices in series (see Section III-B-5).

Figure 21 shows that the cycling characteristics for equivalent MDEIS (with injectors) and OS-SiO$_2$ (without injectors) EAROMs are fairly similar with collapse of the threshold voltage window occurring after approximately the same number of cycles. This is expected if the permanent electron trapping in the SiO$_2$ regions is confined mainly to the intervening OS-SiO$_2$ layer and not to the Si-rich SiO$_2$ injectors[2,5]. Also as expected from previous discussions on I-V characteristics (see Sections III-A-1 and III-B-1, Figs. 6, 11, and 12), the write or erase voltage magnitude for similar charging of the floating gate in terms of V$_T$ (see Eq. 4) is larger and slightly more asymmetric for the device without injectors.

As observed previously with DEIS EAROMs[26], threshold voltage window collapse in MDEIS EAROMs due to permanent electron trapping can be recovered by increasing the write/erase voltage magnitudes. This is shown in Fig. 22 where an MDEIS EAROM is recycled out to 10^{10} cycles with near full threshold voltage window by increasing the write/erase voltage magnitude.

An interesting phenomenon affecting EAROM operation believed to be associated with the physics of the conduction mechanism in OS-SiO$_2$ oxides compared to stoichiometric SiO$_2$ is shown in Figs. 23 and 24. For moderate electric fields, the conduction in OS-SiO$_2$ films is believed to be controlled by direct tunneling between Si islands[20]. The decrease in permanent trapped charge in the OS-SiO$_2$ films for this electric field range is believed, at least in part, to be due to this conduction mechanism since few carriers are capable of being injected into or captured from the SiO$_2$ matrix conduction band[11,20]. At very large electric fields, carrier separation measurements with p-channel FETs[20] show that significant numbers of electrons can get injected into the SiO$_2$ conduction band. This occurs probably via a Fowler-Nordheim mechanism where an electron near the bottom of the conduction band of a Si island tunnels through a triangular barrier into the conduction band of the SiO$_2$ matrix. The critical field at which Fowler-Nordheim injection into the SiO$_2$ becomes important has been shown to increase with increasing Si content and be in the range of average electric fields from $\gtrsim$ 4 to 6 MV/cm[20]. Local electric fields are probably larger due to field enhancement

191

near Si islands and/or due to field screening by the islands which forces more of the applied voltage across the SiO_2 regions. Once the critical field is exceeded and large numbers of electrons can enter the SiO_2 conduction band as "free carriers" and be trapped more readily into energetically deep trapping levels in the SiO_2, a degradation in cycling behavior of EAROMs is expected. This expected degradation with constant transferred charge was seen on the MDEIS devices discussed here and is shown for a particular type of device with a 600 Å thick intervening OS-SiO_2 layer with 3-4% excess atomic Si in Fig. 23. The average electric fields across the OS-SiO_2 layer, as deduced using Eq. 5, for the devices in Fig. 23 are ~4 MV/cm and ~8 MV/cm for 500 μsec and 500 nsec pulsing times, respectively, consistent with the critical field arguments. Even though free electron captured from the SiO_2 conduction band can now take place at very large electric fields, electrons captured in deep traps can still tunnel directly to Si islands under the appropriate electric field conditions. This probably accounts, at least in part, for the observation that although the threshold voltage window starts to collapse sooner at higher fields, it does not collapse as quickly as for a DEIS EAROM at any field (see Figs. 18, 19 and 24). The collapsed threshold voltage window for the MDEIS EAROM which was written and erased with 500 nsec pulses in Fig. 23 could be partially recovered with a 400°C anneal in forming gas for 20 minutes. On recycling after annealing, it collapsed again in a similar fashion. This is analogous to what has been observed for recovering collapsed threshold voltage windows in DEIS EAROMs[3], and it is consistent with thermal electronic discharge of the water-related traps in the SiO_2 matrix[11,12].

Figure 24 shows that for a control DEIS EAROM with a 100 Å intervening SiO_2 layer in the DEIS stack only small differences in the cycling characteristic are seen with similar variations in the write/erase pulse times and fields. DEIS EAROMs with thicker intervening oxide layers show a similar effect on the cycling characteristic with write/erase time under constant threshold voltage window as that shown in Fig. 24. This is expected when the conduction mechanism is strictly controlled by electric field enhanced Fowler-Nordheim tunneling near the Si-rich SiO_2 injector/SiO_2 interface. For short pulsing times ($\lesssim$50 μsec) and high electric fields, some DEIS stacks actually broke down after significant charge build-up in the SiO_2 layer (see Section III-A-2). This is indicated in Fig. 24 when the threshold voltages for the written or erased conditions are equivalent. Also Fig. 24 implies that to first order there is no significant trap creation in the intervening SiO_2 layer of the DEIS stack at the largest average electric fields ($\lesssim$ 10 MV/cm). This is in contrast to high field trap creation reported by others on planar SiO_2 charge transfer devices[27,28] which do not have Si-rich SiO_2 injectors and which would operate at approximately twice the average electric fields used here (assuming planar Fowler-Nordheim tunneling) for similar amounts of

charge transfer to or from the floating gate electrode[3,10,27]. Since MDEIS EAROMs operate at even lower average electric fields than the DEIS EAROMs, it is doubtful that the enhanced threshold voltage window collapse in Fig. 23 can be explained by trap creation.

3. Retention

The level to which the Si content of the intervening oxide layer of the DEIS may be increased is ultimately limited by the degradation in the retention of the device; that is, charge leakage off of or on to the floating gate storage electrode. Figure 25 shows the retention characteristics at room temperature for a grounded control gate condition for the same series of devices as in Fig. 18, after the floating poly-Si layer had been charged from a virgin as-fabricated state with electrons to produce a threshold voltage of $\approx$ +6 to +7 V. Over a 24 hr. period the MDEIS stack, with $\approx$ 3-4% excess atomic Si in the intervening 600 Å thick oxide layer, has the same retention characteristic as the control DEIS structure with the stoichiometric 300 Å thick intervening SiO_2 layer. As seen in Fig. 25, adding more Si to the intervening oxide layer does degrade the retention somewhat. However, it should be pointed out that read disturb effects (that is, electron removal from the floating poly-Si gate via the bottom Si-rich SiO_2 injector caused by the applied positive gate voltage used during the determination of the threshold voltage) are making the retention degradation for devices I and E in Fig. 25 appear worse than actually occurs in a grounded control gate condition. In fact, a large portion of the observed degradation over the first 100 sec on devices I and E is due to the read disturb. In Fig. 26, the retention characteristic in Fig. 25 has been corrected for the read disturb effects. Other DEIS stacks with 1-2% excess atomic Si in a 300 Å thick intervening oxide layer showed the same retention characteristic as the control structure in Fig. 25.

Retention degradations similar to those in Figs. 25 and 26 were seen on devices from the same wafers when electrons were removed from the floating gate, leaving the devices with an initial threshold voltage of $\approx$-1.3 to -1.5 V. This is shown in Fig. 27 where enhanced retention degradation is seen with increasing Si content. Clearly, there is a trade-off with these devices between the maximum number of cycles attainable and long term charge retention due to the added conductivity of the off-stoichiometric intervening oxide layers. Measurements of the retention characteristics on devices similar to those in Figs. 25-27, after at least 10^7 cycles, produced results similar to those in these figures, indicating no pronounced degradation or changes in the materials in the MDEIS or DEIS stacks. Fig. 28 demonstrates this for a device similar to one of those in Fig. 25 with a 600 Å thick off-stoichiometric

193

intervening oxide with 5-6%o excess atomic Si ($R_o=30$), which was cycled out to 4.3×10^9 cycles and had its retention characteristic (after charging the floating poly-Si layer with electrons to a threshold voltage of ≈ 7 V) measured periodically after various amounts of cycling. Observations concerning the retention degradation with cycling similar to those discussed previously for an initial threshold voltage of $\approx +6$ V to $+7$ V were made on similar devices charged to an initial threshold voltage of ≈ -1.3 to -1.5 V. If the device was cycled enough times so that some permanent electron trapping in the SiO_2 regions had occurred, the cycling characteristic actually showed less charge loss or read disturb, at least over times $\lesssim 10^4$ sec (see Fig. 28). This is caused by the repulsive effect of the internal field due to the trapped SiO_2 electrons when the control gate and substrate are grounded. This internal electric field from the trapped electrons in the intervening oxide layer of the MDEIS or DEIS tends to block ejection of electrons off of or injection onto the floating gate by lowering the internal electric field caused by the presence or absence of electrons on the floating gate near the appropriate injecting interfacial region. Degradation phenomenon in retention characteristics has been observed in EAROM structures employing Si_3N_4 layers, usually called metal-nitride-oxide-silicon (MNOS) devices[29].

Figures 29-31 show the effect of temperature between 25°C and 300°C on the retention characteristics for various MDEIS and DEIS EAROMs with stored electrons on the floating gate and with the control gate and substrate at 0 V. For temperatures consistent with normal operating conditions (25°C to 100°C), there is little effect on the retention characteristic. At higher temperatures (>200°C for the structures considered here), there can be considerable charge loss off the floating gate on structures with $OS-SiO_2$ intervening layers containing $\geq 3\%$ excess atomic Si in times less than 10^5 sec (see Figs. 30 and 31). The small amount of electronic charge loss off the floating gate of the DEIS EAROM (see Fig. 29) at elevated temperatures is consistent with previously reported results[3]. Figure 32 shows the effect of temperature on the retention characteristic of an MDEIS EAROM with an $R_o=40$ (3-4% excess atomic Si) $OS-SiO_2$ intervening oxide in the MDEIS stack under grounded control gate and substrate conditions which has the floating gate in an initial positively charged state. The temperature dependence under $V_g=0$ V conditions for positive or negative stored charge on the floating gate are similar (compare Figs. 30 and 32).

The effect of control gate bias on the retention characteristics of MDEIS EAROMs for various temperatures and charge states of the floating gate are shown in Figs. 33-36. For stored electrons on the floating gate electrode, increasing positive voltages on the control gate will accelerate charge loss by pulling electrons back to the control gate due to the non-ohmic conductivity of the $OS-SiO_2$ in the MDEIS stack. Negative voltages on the control gate will

opposed the internal field of the stored electrons on the floating gate and minimize their loss (see Fig. 33). As shown in Fig. 34, the opposite is true for stored positive charges (ionized donors) on the floating gate where negative control gate voltages increase compensation of the positive charges by electron injection from the control gate while positive voltages oppose the internal electric field of the positive charges. Figs. 35 and 36 show that for a constant applied control gate voltage, field induced negative charge loss (Fig. 35) or positive charge loss (Fig. 36) off the floating gate is accelerated significantly with increasing temperature (particularly for temperatures $\geq$ 200°C) on MDEIS EAROMs with OS-SiO_2 intervening oxide layers with >3% excess atomic Si.

The dependence of the retention characteristics of MDEIS EAROMs on control gate voltage and temperature should be predictable from current density-electric field relationships of MDEIS or OS-SiO_2 layers in capacitor configurations like that shown in Fig. 8. However, the charge integrating ability of the floating gate EAROM devices allows us to calculate the <u>average</u> current-field relationship from retention characteristics directly. In fact, because of this charge integrating ability of EAROMs, currents $\leq 10^{-15}$ A at electric fields $\leq$ 1 MV/cm can be easily measured which is difficult, if not impossible, using direct current measuring techniques on capacitors structures (as described in Section III-A-1). Figure 37 shows the magnitude of the average current density as a function of temperature for various field conditions deduced from the data of Figs. 30, 32, 35 and 36. The average particle current $J_{p_{ave}} = -C_1^* \Delta[\Delta V_T]/A_1 t$ was deduced using Eq. 4 and the average electric field across the intervening oxide layer of MDEIS E_{o_1} is given by Eq. 5. Since the calculations were done for small changes in $\Delta[\Delta V_T]$, E_{o_1} is approximately constant and indicated by the values in Fig. 37. These data show a two to three order of magnitude change in current with temperature from 25°C to 300°C for fields < 1 MV/cm similar to the low field data from large area capacitors (see Fig. 8).

DEIS EAROMs with stoichiometric intervening SiO_2 in the DEIS stack have charge loss at low internal electric fields and high temperatures (>200°C, see Fig. 29). This charge loss is <u>not</u> limited by the injection of electrons via Fowler-Nordheim tunneling from near the Si-rich SiO_2 injector/SiO_2 interfaces as observed at larger fields[3]. It is, however, limited by thermal activation over ~ 3 eV energy barriers near the n-degenerate poly-Si floating or control gates for negative charge loss or for positive charge loss by electron compensation, respectively [3]. Thermal activation over energy barriers (which are ~ 3 eV [3,6,9,11]) created by the Si islands in the OS-SiO_2 films of the MDEIS can also be a contributing mechanism for loss of floating gate charge on MDEIS EAROMs with low internal electric fields (see Figs. 30 and 31) and high temperatures (>200°C). These energy barriers might be lowered by tenths

of an electron-volt (by a Poole-Frenkel or internal Schottky effect) if local fields near the Si islands are very large because of island curvature and/or field screening by Si islands forcing more of the applied voltage across the SiO_2 matrix portion of these materials [20,25,30]. Furthermore, thermally-assisted tunneling phenomena between the Si islands in the $OS-SiO_2$ films may also be operative in controlling current-field characteristics and therefore charge loss off the floating gate at low fields and high temperatures [6,31].

Retention characteristics similar to those presented in Figs. 25-36 for similar experimental conditions were seen on several wafer runs besides XDEIS II 5 with, in some cases, somewhat different processing. For example, devices like those in Fig. 20 with equal floating and control gate areas, requiring one less masking step, gave similar retention results as a comparable XDEIS II 5 wafer (see Fig. 31, wafer I) at similar temperature and electric field conditions of the MDEIS stack.

4. Gate SiO_2 - Charge Trapping and Interface States

Charge trapping in the thermally grown gate oxide between the Si substrate and the floating poly-Si gate and surface state creation at the gate oxide-substrate Si interface was neither expected nor observed for the EAROM devices. Even for 10^{11} write/erase cycles with equal or unequal control and floating gate areas (see Fig. 1) much less than 10^{-3} $coul/cm^2$ total electronic charge will pass through the gate oxide. For significant charge trapping in the gate oxide, $\geq 10^{-3}$ $coul/cm^2$ total charge must be transferred through the SiO_2 layer for normal amounts of electron trapping centers with densities of $\geq 10^{12}$ cm^{-2} and capture cross sections of $\leq 10^{-17}$ cm^2 [5,7,11-15]. Surface state generation as measured using capacitance-voltage techniques on large, equal area control gate and floating gate devices was not observed for $\geq 10^8$ write/erase cycles on structures similar to those described here. Also no distortions in the device drain current vs. gate voltage curves were observed on the same devices for the same cycling conditions. By decreasing the control gate area with respect to the floating gate area (as was done on most of the devices described here, in particular the XDEIS II 5 wafers), even smaller electric fields and therefore smaller leakage currents will appear in the gate oxides of the devices[4,5].

5. Other Device Configurations

Although the three port (source, drain, and control gate) devices described here with DEIS or MDEIS stacks placed inbetween the control and floating gates need both negative and positive voltages applied to the control gate in order to be written and erased, other device geometries would require only positive voltages[4,5]. For example, the addition of

another larger coupling gate to a structure with the DEIS or MDEIS inbetween the floating gate and a smaller control gate would allow writing by grounding the smaller control gate and biasing the larger coupling gate at a positive voltage. Erasing would be obtained by grounding the coupling gate and biasing the control gate at a positive voltage. Another example would be for a case where the DEIS or MDEIS stack is placed inbetween a diffusion in the single crystal Si substrate and the floating gate electrode with an SiO_2 layer now forming the dielectric inbetween the floating gate and the control gate electrode. This device could be written by applying a positive voltage to the control gate with the diffusion grounded and could be erased by applying a positive voltage to the diffusion with the control gate grounded.

To minimize read disturb effects in MDEIS EAROMS, a two device cell with a MOS FET in series with the MDEIS EAROM could be used with some lost in area. Such cells are typical of commercially available memory chips, particularly MNOS EAROMs[32,33]. In this cell once the EAROM is written or erased, its control gate would be be grounded. The read operation would be performed by addressing the series MOS FET and turning on its conducting channel with positive gate voltage in the case of an n-channel FET. If the MDEIS is erased (positive charge on the floating gate), its channel will be turned on (assuming it is an n-channel device), and current will flow through the series of FETs to ground. If the MDEIS is written (negative charge on the floating gate), its channel will be turned off, and current will not flow through the series of FETs. Therefore the two device cell performs a non-volatile memory function having two information states that can be sensed by current flow through the FET channels. Since the control gate of the MDEIS EAROM is always kept at or near ground after information is stored during a write or erase cycle, read disturb effects caused by directly applying positive read voltages to the control gate of the MDEIS EAROM will not occur.

IV Modeling

Previous modeling calculations which were used to predict values of $\Delta[\Delta V_T]$ for write/erase times from 1 μsec to 5 msec on DEIS EAROMs used a simple exponential approximation to the Fowler-Nordheim current-voltage relationship[4]. This exponential relationship for the particle current per unit area J_p flowing between the control and floating gates was given by $J_p = J_p^i \exp\left[S(E_{o_1} - E_{o_1}^i)\right]$ where J_p^i, $E_{o_1}^i$ and S are constants[4]. However for OS-SiO_2 (no injectors) or MDEIS (with injectors) EAROMs this exponential approximation does not give a reasonable fit to the $\Delta[\Delta V_T)]$ data for write/erase times ≤ 50 μsec. A closed form relationship for $\Delta[\Delta V_T]$ or ΔV_T can be derived using Eq. 1 which assumes a Fowler-Nordheim like J_p vs. E_{o_1} condition for the particle current and Eq. 5 which gives E_{o_1} in terms of the gate voltage and relevant device capacitances. As was discussed in Section

III-A-1, the pre-exponential term of the J_p vs. E_{o_1} relationship is not critical as long as the $\exp[-b/E_{o_1}]$ term dominates. Therefore, $J_p = aE_{o_1}^2 \exp[-b/E_{o_1}]$ will be used here. Also b is not strictly a constant, but can show a field dependence[20]. However, to first order this will be neglected as will the subtle space charge effects discussed in Section III-B-1. Since $J_p = (dQ/dt)/A_1$ [3,4], it can be shown using Eq. 4 that

$$J_p = -C_1^*(dV_T/dt)/A_1 \tag{6}$$

Using the Fowler-Nordheim relationship $J_p = aE_{o_1}^2 \exp[-b/E_{o_1}]$ in Eq. 6, one obtains

$$\frac{dV_T}{dt} = -(A_1aE_{o_1}^2/C_1^*) \exp[-b/E_{o_1}] \tag{7}$$

Substituting Eq. 5 for E_{o_1} in Eq. 7, rearranging terms, and integrating both sides from 0 to t yields the closed form relationship for $\Delta V_T(t)$ [34].

$$\Delta V_T = \frac{-\chi(V_g - \Phi_{ms} - \Psi_s)}{C_1^*} + \tag{8}$$

$$\frac{C_T \ell_{o_1} b / C_1^*}{\ln\left[\dfrac{A_1 abt}{\ell_{o_1}C_T} + \exp\left\{\dfrac{C_T \ell_{o_1} b}{\chi(V_g - \Phi_{ms} - \Psi_s) + C_1^* \Delta V_T(0)}\right\}\right]}$$

In the limits of $t \to 0$ or $t \to \infty$, Eq. (8) reduces to $\Delta V_T = \Delta V_T(0)$ or $\Delta V_T = -\chi(V_g - \Phi_{ms} - \Psi_s)/C_1^*$ respectively. Figure 38 shows a comparison between $\Delta[\Delta V_T]$ vs. V_g data and Eq. 8 for writing from a $V_T(0) = 2.6$ V state on an MDEIS EAROM. The Fowler-Nordheim constants a and b were determined from least squares data fitting of ramped ($|dV_g/dt| = 0.5$ V/sec) I-V characteristics obtained from capacitors of .005 cm^2 area fabricated simultaneously with the EAROM devices. As seen in this figure, the predicted results are in good agreement with the measured data at all write gate voltages except where $\Delta[\Delta V_T]$ becomes approximately constant. This occurs when the floating gate becomes over charged and electrons leak back to the control gate (as discussed in Section III-B-1).

V. Conclusions

The data presented on MDEIS EAROMS in this study demonstrates that the quest for the elusive "pure" non-volatile random access memory may be ending. OS-SiO$_2$ films with or without Si-rich SiO$_2$ injectors used in EAROM type devices give low average electric field

(and therefore power) operation with a least four orders of magnitude improvement in cyclability over structures using stoichiometric SiO_2. With some of the devices described here, $> 10^{10}$ cycles have been achieved without total collapse of the threshold voltage window due to electron trapping in the SiO_2 regions of the OS-SiO_2 films. A modification of the conduction mechanism by the presence of the tiny Si islands in the OS-SiO_2 films in which most carriers (electrons) move from Si island to Si island by direct tunneling rather than through the SiO_2 matrix conduction band is believed to be, at least in part, responsible for some of the observed improvements over DEIS EAROMs. Further decreases in permanent electronic charge trapping are also believed to be due to the "stepping stone" nature of the Si islands to which electrons trapped on energetically deep water-related impurity sites in the SiO_2 matrix portion of the OS-SiO_2 films can directly tunnel under appropriate field conditions.

Although some read disturb effects were seen on certain "single-cell" devices, charge retention of the floating gate was not effected drastically for the devices studied here for a grounded gate condition at normal operating temperatures between 25-100°C. For temperatures $\geq$ 200°C, significant charge loss was observed on devices with thin OS-SiO_2 layers with the highest percentage of excess atomic Si studied here (5-6%). This enhanced charge loss is believed to be due to the bulk-limited low-field conduction mechanism which is controlled by thermally assisted tunneling (percolation).

Minimization of low electric field breakdowns was observed due to the field screening ability of these films, particularly for contact inhomogenities which are believed to be a leading cause of this breakdown phenomenon. These OS-SiO_2 films in MDEIS EAROMs could also be made thicker than SiO_2 films in DEIS EAROMs due to their enhanced conductivity and lower charge trapping which adds more flexibility to making these layers reproducibly with CVD systems. Also thicker films minimize possible pinhole problems which might occur in manufacturing.

The EAROM device operation was shown to be predictable (at least, to first order) from ramped current-voltage data obtained from large area capacitors and a simple physical model based on quantum mechanical tunneling. At very large average electric fields, some increased degradation in the cyclability was seen. It was proposed that this degradation was due to enhanced charge trapping which occurred when large numbers of electrons could tunnel into and be captured more effectively from the conduction band of the SiO_2 matrix portion of the OS-SiO_2 films.

199

Future directions with MDEIS EAROMs in two device cells using an FET transfer gate should maximize cyclability and low voltage operation while at the same time minimizing read disturb. All positive voltage operation from 5 V supplies should be possible for functional EAROM or NVRAM chips using these materials.

Acknowledgements

References

1. D.J. DiMaria, D.W. Dong, C. Falcony, and S.D. Brorson, IEEE Electron Dev. Lett. EDL-3, 191 (1982).

2. D.J. DiMaria, K.M. DeMeyer, and D.W. Dong, IEEE Elecron Dev. Lett. EDL-1, 179 (1980).

3. D.J. DiMaria, K.M. DeMeyer, C.M. Serrano, and D.W. Dong, J. Appl. Phys. 52, 4825 (1981).

4. D.J. DiMaria, K.M. DeMeyer, and D.W. Dong, IEEE Trans. Electron Dev. ED-28, 1047 (1981).

5. C. Falcony, D.J. DiMaria, D.W. Dong, and K.M. DeMeyer, J. Appl Phys. 53, 43 (1982).

6. D.J. DiMaria and D.W. Dong, J. Appl. Phys. 51, 2722 (1980).

7. D.J. DiMaria, R. Ghez, and D.W. Dong, J. Appl. Phys. 51, 4830 (1980).

8. D.W. Dong, E.A. Irene, and D.R. Young, J. Electrochem. Soc. 125, 810 (1978).

9. D.J. DiMaria, in The Physics of MOS Interfaces, edited by G. Lucovsky, S.T. Pantelides, and F.L. Galeener (Pergamon, New York, 1980) p. 1, and references contained therein.

10. D.J. DiMaria and D.W. Dong, Appl. Phys. Lett. 37, 61 (1980).

11. D.J. DiMaria, in the The Physics of SiO_2 and Its Interfaces, edited by S.T. Pantelides (Pergamon, New York 1978), p.160, and references contained therein.

12. E.H. Nicollian, C.N. Berglund, P.F. Schmidt, and J.M. Andrews, J. Appl. Phys. 42, 5654 (1971).

13. F.J. Feigl, D.R. Young, D.J. DiMaria, S.K. Lai, and J.A. Calise, J. Appl. Phys. 52, 5665 (1981).

14. A. Hartstein and D.R. Young, Appl. Phys. Lett. 38, 631 (1981).

15. D.R. Young, E.A. Irene, D.J. DiMaria, and R.F. DeKeersmaecker, J. Appl Phys. 50, 6366 (1979).

16. D.D. Rathman, F.J. Feigl, and S.R. Butler, in Insulating Films on Semiconductors, edited by G. Roberts (London: Institute of Physics and Physical Society, 1980), p. 48.

17. S.K. Lai, D.R. Young, J.A. Calise, and F.J. Feigl, J. Appl. Phys. 52, 5691 (1981).

18. D.R. Young, J. Appl Phys. 52, 4090 (1981).

19   D.J. DiMaria, D.W. Dong, C. Falcony, S.K. Lai, unpublished.

20. D.J. DiMaria, D.W. Dong, C. Falcony, T.N. Theis, J.R. Kirtley, J.C. Tsang, D.R. Young, F.L. Pesavento, and S.D. Brorson, J. Appl. Phys., Ms. #R-3726.

21. M. Lenzlinger and E.H. Snow, J. Appl. Phys. 40, 278 (1969).

22. B. Abeles, P. Sheng, M.D. Coutts and Y. Arie, in Advances in Physics, edited by D.H. Martin, B.R. Coles, and N.F. Mott (Taylor and Francis Ltd., London, 1975), p.p. 407-461, Vol 24.

23. P. Solomon, J. Vac. Sci. Technol. 14, 1122 (1977).

24. E. Harari, J. Appl. Phys. 49, 2478 (1978).

25. S.M. Sze, Physics of Semiconductor Devices, (Wiley-Interscience, New York, 1981), Chaps. 7 and 8.

26. D.J. DiMaria, in Insulating Films on Semiconductors), edited by M. Schulz and G. Pensl (Springer-Verlag, Berlin, 1981), p.88.

27. C.S. Jeng, T.R. Ranganath, C.H. Huang, H.S. Jones, and T.T.L. Chang, in IEDM Technical Digest 1981, (IEEE, Inc., New York, 1981), p.388.

28. A. Badihi, B. Eitan, I. Cohen, and J. Shappir, Appl. Phys Lett. 40, 396 (1982).

29. M.H. White, J.W. Dzimianski, and M.C. Peekerar, IEEE Trans. on Electron Devices ED-24, 577 (1977).

30. R.M. Hill, Philos. Mag. 23, 59 (1971).

31.   V. Ambegaokar, B.I. Halperin, and J.S. Langer, Phys. Rev. $\underline{B4}$, 2612 (1971).

32.   R.P. Capece, Electronics $\underline{52}$, 39 (1979).

33.   B. Giebel, IEEE J. Solid-St. Circuits $\underline{SC-15}$, 311 (1980).

34.   S.T. Wang, IEEE Trans. Electron Dev. $\underline{ED-27}$, 297 (1980).
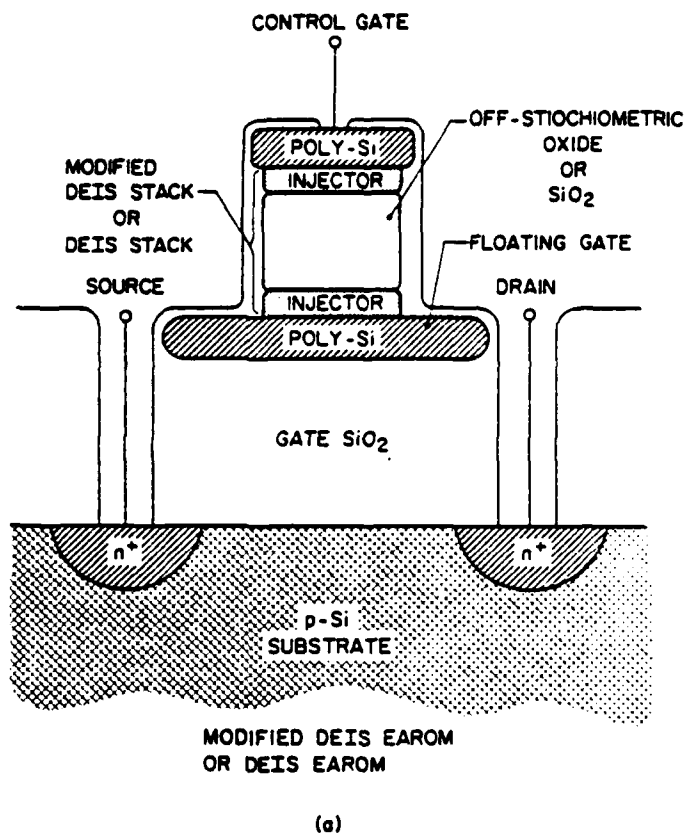
(a)

Fig. 1   Schematic representation of a three port DEIS or MDEIS EAROM with an SiO_2 or off-stoichiometric oxide intervening layer in the DEIS OR MDEIS stack, respectively. The control (top) gate and floating (bottom) gate are formed from n-degenerate poly-Si. The DEIS or MDEIS stack is incorporated inbetween the control and floating gates. A thermal oxide layer is grown from the Si substrate and used as a gate insulator inbetween the floating gate and the Si substrate. The source and drain contacts complete the FET portion of the structure. Not drawn to scale.
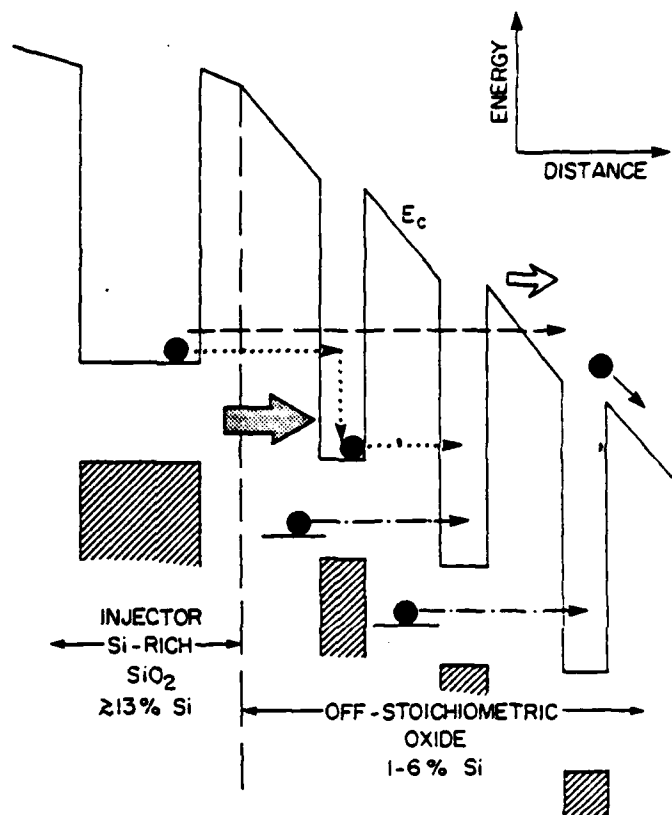
Fig. 2   Schematic energy band representation of the modification in the conduction mechanism when the SiO_2 layer of a DEIS stack is replaced by an OS-SiO_2 layer. The OS-SiO_2 layer (1-6% excess atomic Si) is thought to perturb the SiO_2 bands as shown because of the presence of small Si inclusions similar to the Si-rich SiO_2 injector layers which have $\geq$ 13% excess atomic Si. The dashed arrow shows normal Fowler-Nordheim tunnel injection from a Si island conduction band to the SiO_2 conduction band occurring in a DEIS at the Si-rich SiO_2/SiO_2 interface, while the dotted arrow shows the injection and conduction from Si island conduction band to Si island conduction band which is thought to occur in the MDEIS. Fowler-Nordheim tunneling from an injector into the SiO_2 conduction band can take place in an MDEIS in regions where the OS-SiO_2 does not have a Si island within 30-50 Å of the interface. The dot-dashed line represents possible tunneling of electrons from "deep" traps in the SiO_2 matrix of the OS-SiO_2 film to the near-by conduction band of a Si island. The magnitude of the large arrows in the figure indicates that larger currents for the same average fields are observed in MDEIS as compared to DEIS stacks, at least for times $\geq$ 500 μsec.
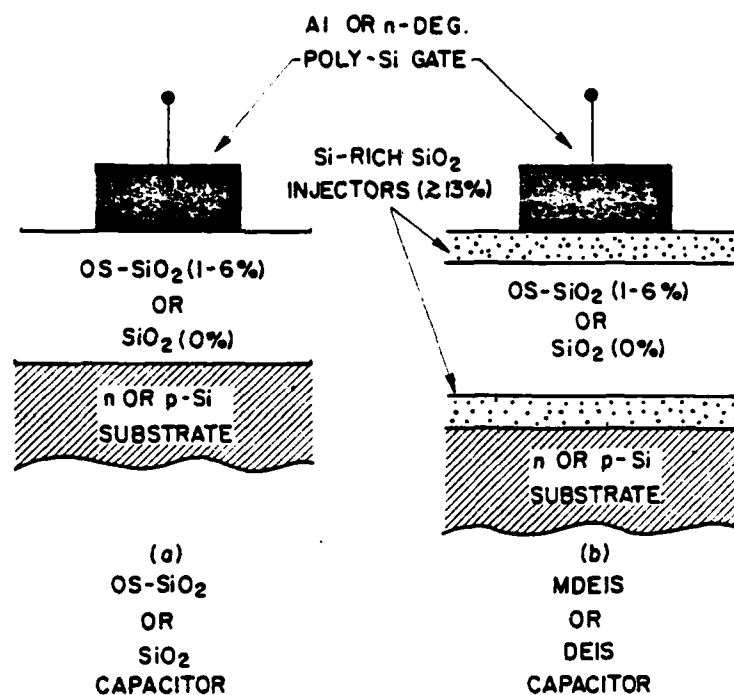
Fig. 3   Schematic representation of the types of capacitor structures used for many of the measurements discussed in this paper. (a) Simple MOS capacitor with either an SiO₂ or off-stoichiometric oxide sandwiched between the metal (A1) or n-degenerate poly-Si gate electrode and the Si substrate (n or p-type <100>); (b) similar to (a) only with a DEIS (an intervening SiO₂ and two Si-rich SiO₂ injectors) or MDEIS (an intervening off-stoichiometric oxide and two Si-rich SiO₂ injectors) sandwiched between metal or poly-Si gate electrodes and the Si substrate.
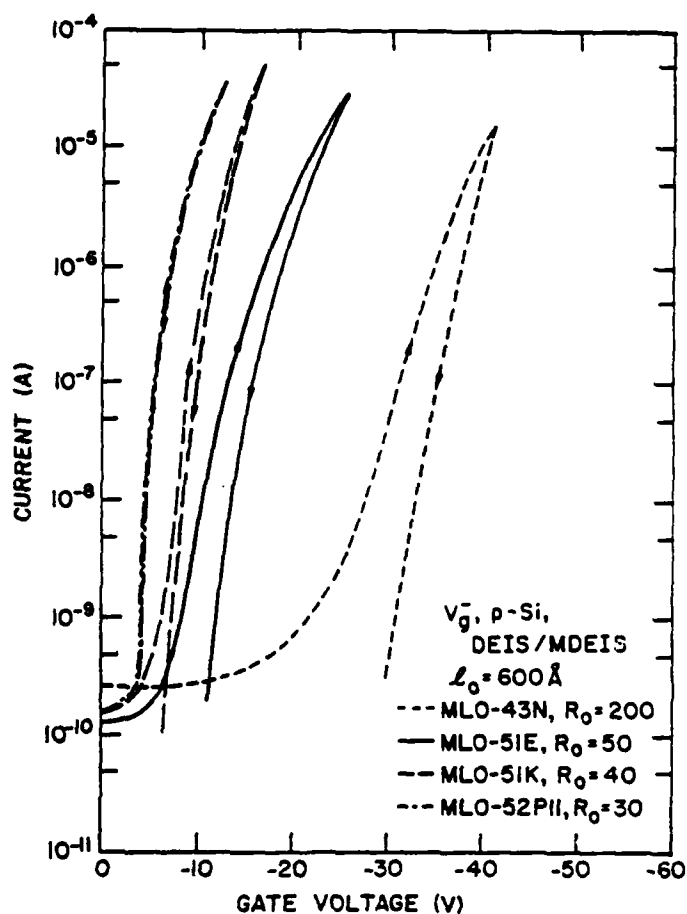
Fig. 4   Magnitude of the dark current at room temperature as a function of ramped gate voltage ($| \, dV_g/dt \, | = 0.5$ V/sec) for SiO_2 and off-stoichiometric oxides with 1-2% ($R_o = 50$) to 5-6% ($R_o = 30$) excess atomic silicon in DEIS OR MDEIS stacks (see Fig. 3b), respectively. The oxide layers were 600 Å thick, the Si-rich SiO_2 injectors ($R_o = 3$) were each 200 Å thick, and the metal gate electrode was Al. A virgin capacitor was used from each wafer. As the excess silicon content increases in the oxide, the I-V curves shift towards lower gate voltages and the hysterisis associated with permanent charge trapping diminishes.
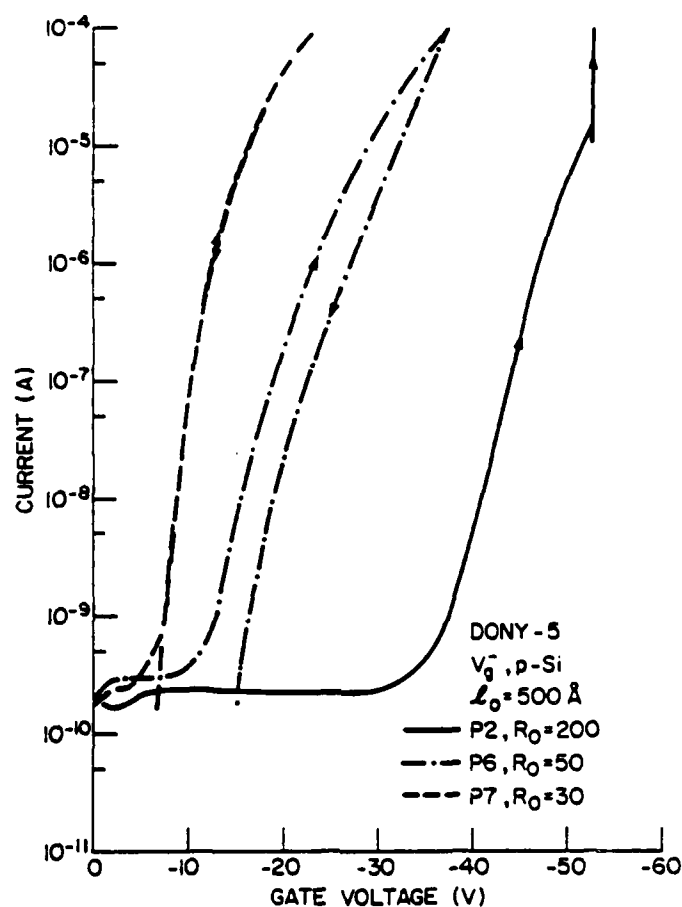
Fig. 5   Magnitude of the dark current at room temperature as a function of ramped gate

voltage (| dV_g/dt | $=$ 0.5 V/sec) for SiO_2 and off-stoichiometric oxides similar to

those in Fig. 4 only with the oxide in an MOS configuration without the Si-rich

SiO_2 injectors (see Fig. 3a). A virgin capacitor was used from each wafer. The

behavior observed is similar to that observed in Fig. 4 except that the stoichiome-

tric SiO_2 MOS capacitor could not be ramped to higher than $\approx 10^{-5}$ A without the

oxide breaking down destructively.
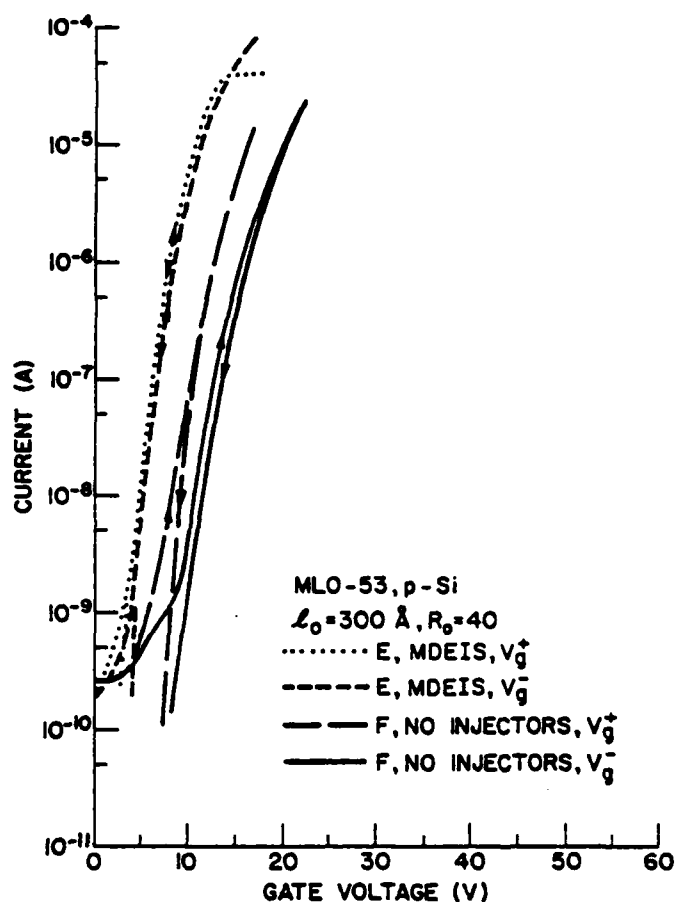
Fig. 6   Magnitude of the dark current at room temperature as a function of the magnitude of the ramped gate voltage ($|$ dV_g/dt $|$ $=$ 0.5 V/sec) for capacitors using off-stoichiometric oxides (R_o=40) with (see Fig 3b) and without (see Fig. 3a) Si-rich SiO_2 (R_o=3) injectors. The oxide layers were 300 Å thick, the Si-rich SiO_2 injectors were each 200 Å thick, and the gate electrode was Al. Both polarities are shown with a virgin capacitor ramped for each polarity.
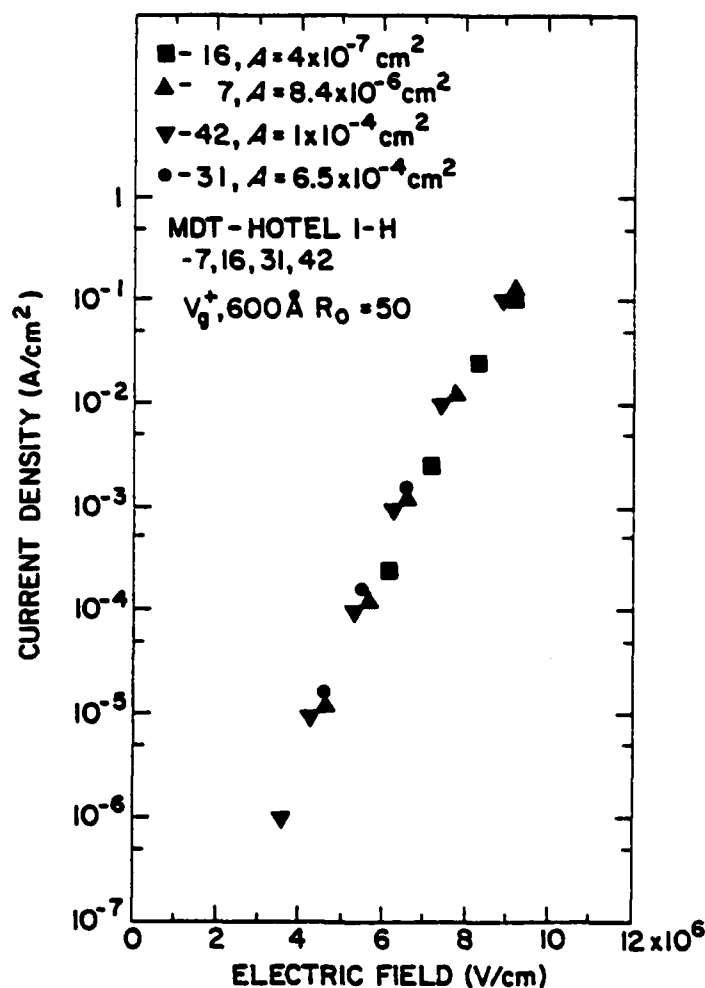
Fig. 7   Magnitude of the areal current density as a function of the magnitude of the
average electric field for various gate areas A on a 600 Å thick OS-SiO$_2$ layer with
1-2% (R_o=50) excess atomic Si in an MOS capacitor configuration (see Fig. 3a).
The current was obtained on a virgin capacitor for each area indicated at a positive
gate polarity and a voltage ramp rate magnitude of 0.5 V/sec. The gate electrode
was n-degenerate poly-Si. This figure indicates the very weak dependence of the
dark currents on gate area or periphery for OS-SiO$_2$ layers.
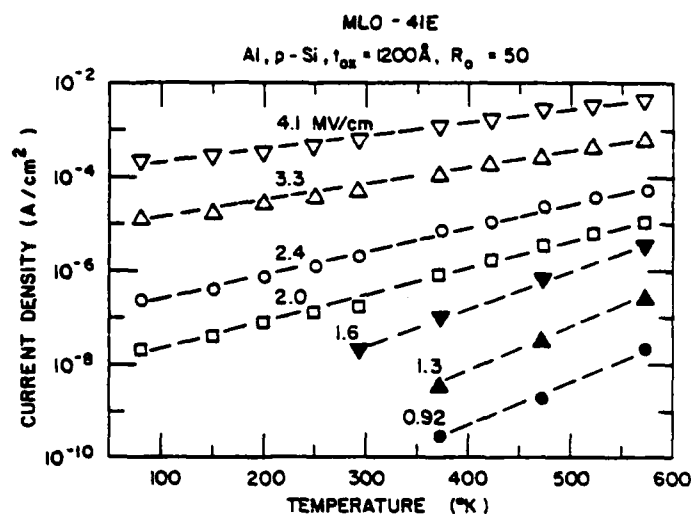
MLO - 4IE

Al, p - Si, t_{ox} = I200Å, R_o = 50

Fig. 8   Magnitude of the dark current as a function of temperature for several negative

gate voltages on a 1200 Å thick OS-SiO$_2$ layer (R_o=50, 1-2% excess atomic Si)

incorporated into an MOS capacitor configuration with an Al gate electrode (see

Fig. 3a).   Data for gate voltages between -12 V and -20 V were obtained from

point by point measurements, while data for gate voltages from -25 V to -50 V

were obtained from ramp I-V measurements.   This data shows that the temperature

dependence decreases with increasing electric field.
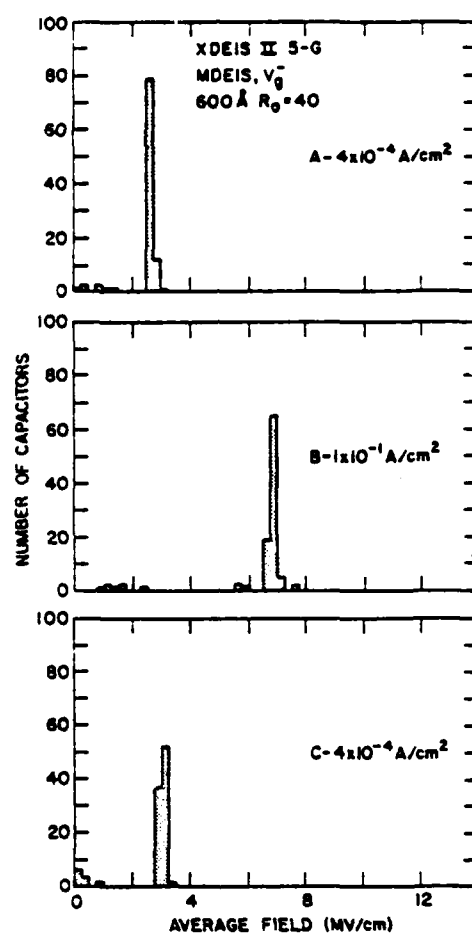
Fig. 9   Histogram in 0.25 MV/cm bins of the number of capacitors (0.005 cm² area) on
          an MDEIS wafer (XDEIS II 5-G) to draw a current of 4×10^{-4} A/cm² or 1×10^{-1}
          A/cm² as a function of the magnitude of the average field in the 600 Å thick
          OS-SiO$_2$ layer with 3-4% excess atomic Si (R_o=40). Si-rich SiO$_2$ injectors were
          200 Å thick and all samples were ramped with negative gate voltage bias from 0 V
          at a rate of -5 V/sec. An initially as-fabricated virgin wafer was put through the
          sequence shown by figures A, B, and C in which the same 100 capacitors were
          sequentially ramped to current levels of A - 4×10^{-4}A/cm², B - 1×10^{-1} A/cm², and
          C - 4×10^{-4} A/cm².
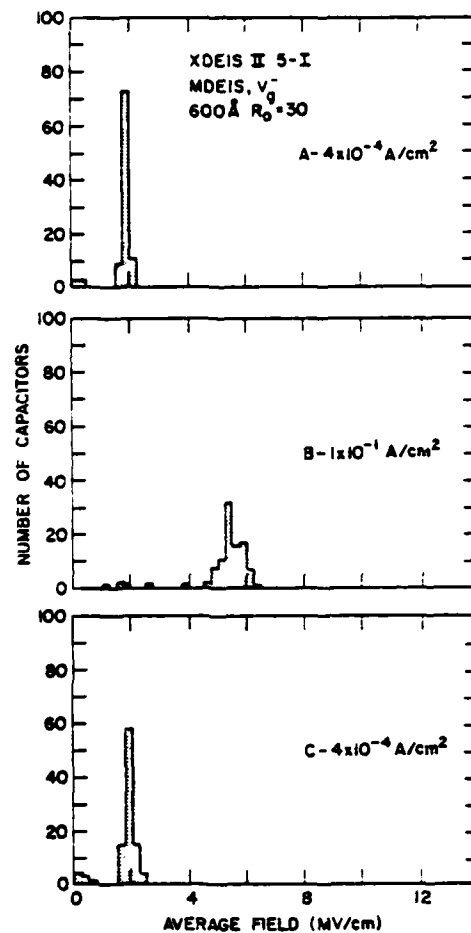
Fig. 10   Histogram on an MDEIS wafer (XDEIS II 5-I) as a function of the magnitude of the average field. Experimental conditions and sequence are the same as in Fig. 9. MDEIS wafer XDEIS II 5-I is similar to the one used in Fig. 9 (XDEIS II 5-G) except that the OS-SiO$_2$ layer contained 5-6% excess atomic Si (R$_o$=30).
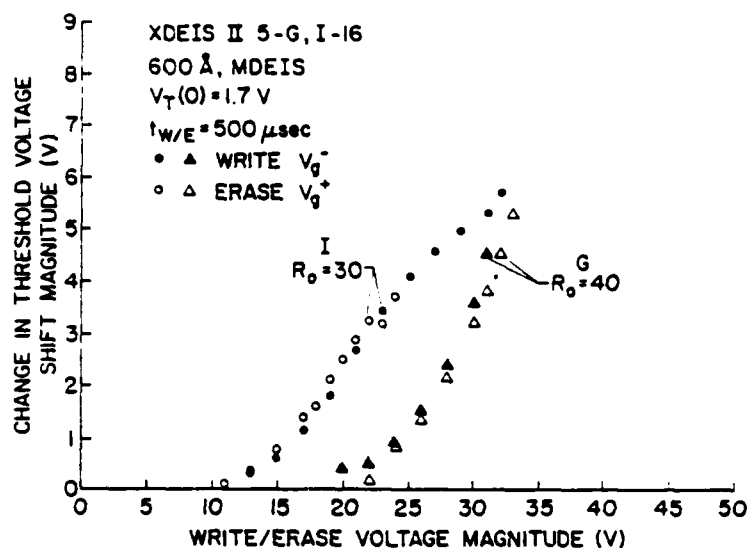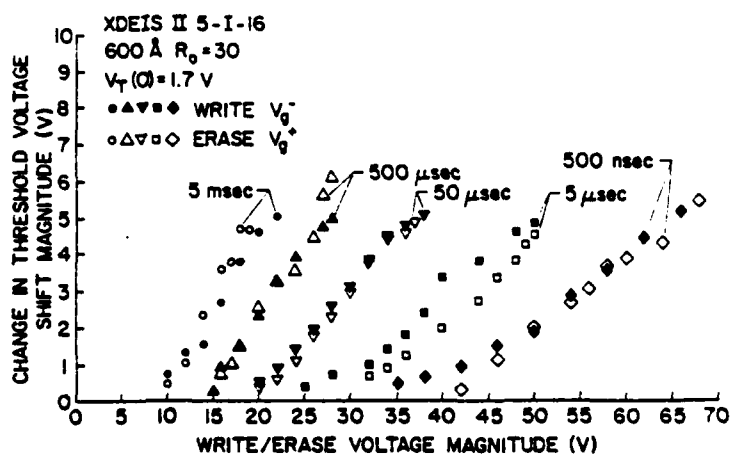
Fig. 11   Change in the threshold voltage shift magnitude as a function of write/erase voltage magnitude on MDEIS EAROMs (XDEIS II 5-G, I-16) with 3-4% (R_o=40) and 5-6% (R_o=30) excess atomic Si, respectively, in the 600 Å thick intervening OS-SiO_2 layer of the MDEIS stack (see Fig. 1). The Si-rich SiO_2 injectors were 200 Å thick, and the gate oxide was 650 Å thick. The initial threshold voltage was 1.7 V, and the write/erase pulse time was 500 μsec. This figure shows that as the Si content in the OS-SiO_2 layer increases the device can be written or erased at lower gate voltages for times $\geq$ 500 μsec.
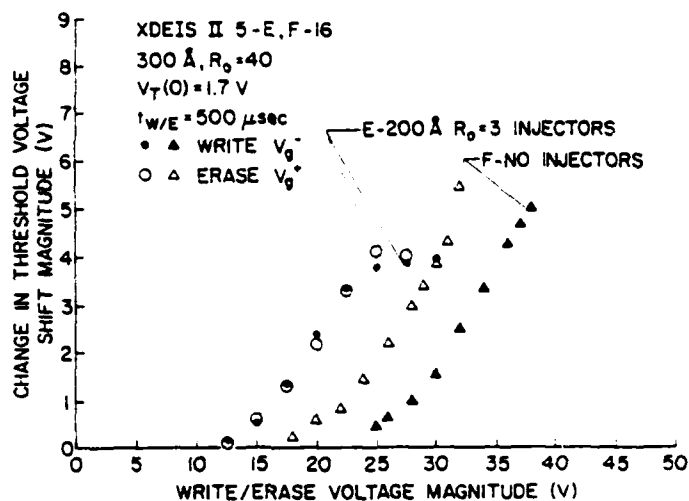
Fig. 12. Change in the threshold voltage shift magnitude as a function of write/erase voltage magnitude on EAROMs (XDEIS II 5-E, F-16) with (E) and without (F) 200 Å thick Si-rich SiO_2 injectors. The OS-SiO_2 layer of both types of devices was 300 Å thick and contained 3-4% excess atomic Si ($R_o = 40$). The gate oxide was 650 Å thick. Operating conditions are similar to Fig. 11. This figure shows that EAROM devices with injectors can be written or erased at lower gate voltages than *structures without injectors.*
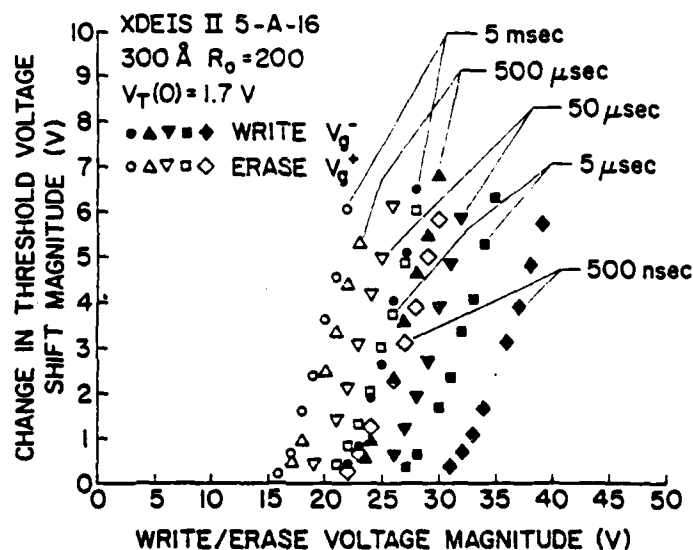
Fig. 13 Change in the threshold voltage shift magnitude as a function of write/erase voltage magnitude on a *DEIS EAROM* (*XDEIS II* 5-A-16) with a 300 Å thick intervening SiO_2 ($R_o = 200$) layer in the DEIS stack (see Fig. 1) for various write/erase pulse times from 5 msec to 500 nsec. The Si-rich SiO_2 injectors were 200 Å thick, and the gate oxide was 650 Å thick. The initial threshold voltage was 1.7 V.

215
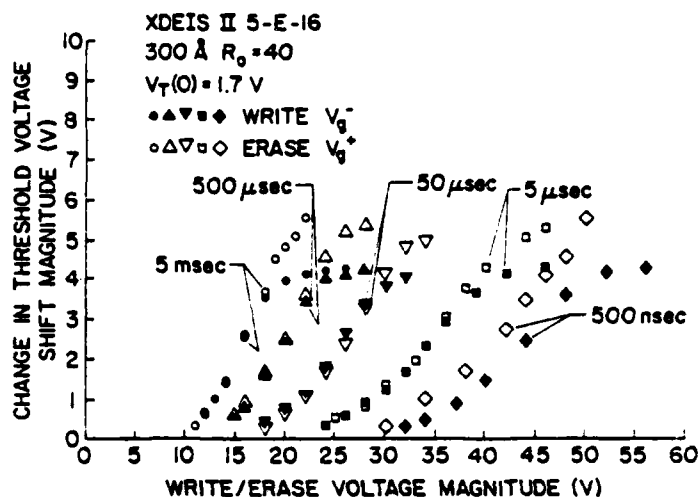
Fig. 14   Change in the threshold voltage shift magnitude as a function of write/erase voltage magnitude on a MDEIS EAROM (XDEIS II 5-E-16) with a 300 Å thick intervening OS-SiO$_2$ layer with 3-4% excess atomic Si (R$_0$=40) in the MDEIS stack (see Fig. 1) for various write/erase pulse times from 5 msec to 500 nsec. The Si-rich SiO$_2$ injectors were 200 Å thick, and the gate oxide was 650 Å thick. The initial threshold voltage was 1.7 V.

Fig. 15   Change in the threshold voltage shift magnitude as a function of write/erase voltage magnitude on a MDEIS EAROM (XDEIS II 5-I-16) with a 600 Å thick intervening OS-SiO$_2$ layer with 5-6% excess atomic Si (R$_0$=30) in the MDEIS stack (see Fig. 1) for various write/erase pulse times from 5 msec to 500 nsec. The Si-rich SiO$_2$ injectors were 200 Å thick, and the gate oxide was 650 Å thick. The initial threshold voltage was 1.7 V.
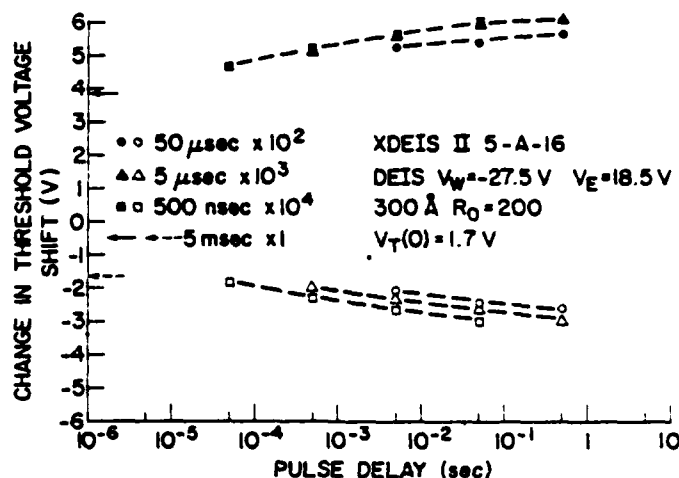
Fig. 16   Change in the threshold voltage shift as a function of the time delay between write or erase pulses for pulse trains of various pulse time duration but all adding to 5 msec on a DEIS EAROM similar to that described in the caption of Fig. 13. Each point was taken by either writing with a voltage of -27.5 V (as indicted by solid symbols) or erasing with a voltage of 18.5 V (as indicated by open symbols) using the number of pulses with the indicated pulse time duration listed opposite the appropriate symbols. Prior to each write or erase operation the device was reset to an initial threshold voltage $V_T(0) = 1.7$ V. These data show that for a DEIS EAROM a sum of pulses of a given total time duration can give a larger change in threshold voltage shift magnitude than a single pulse.
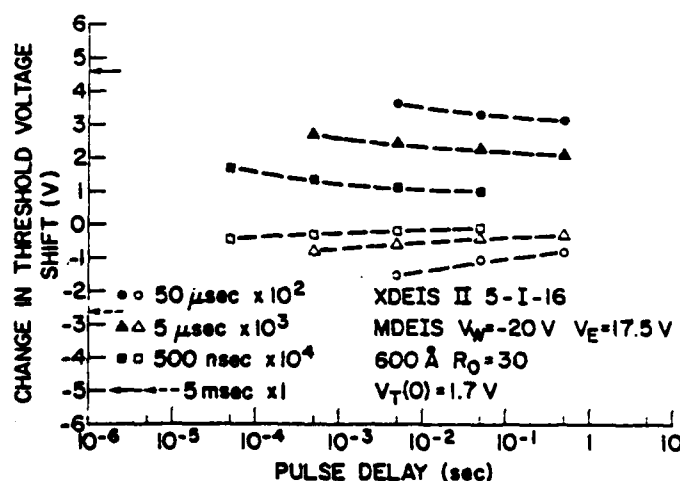
Fig. 17   Change in the threshold voltage shift as a function of the time delay between write or erase pulses for pulse trains of various pulse time duration but all adding to 5 msec on a MDEIS EAROM similar to that described in the caption of Fig. 11. Symbols have the same meanings as in Fig. 16, but with write or erase voltages of -20 V or 17.5 V, respectively. The trend in these data is opposite that of Fig. 16 where a stoichiometric oxide layer was used in the DEIS stack.
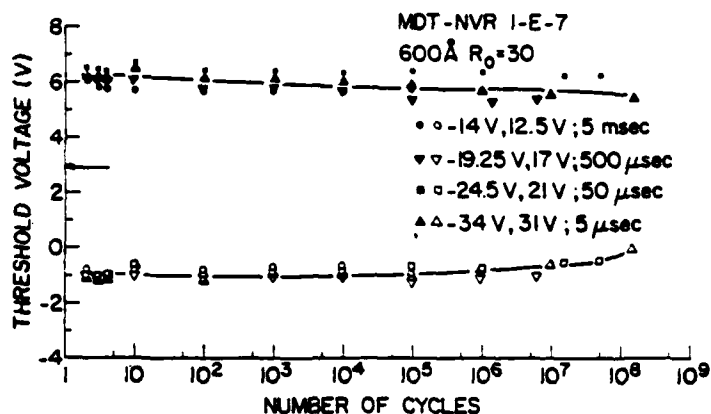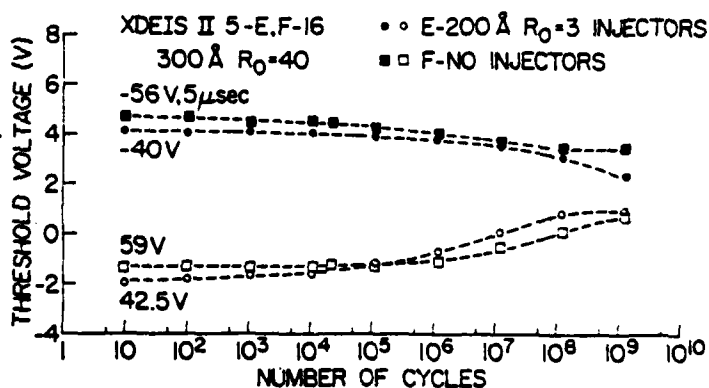
217
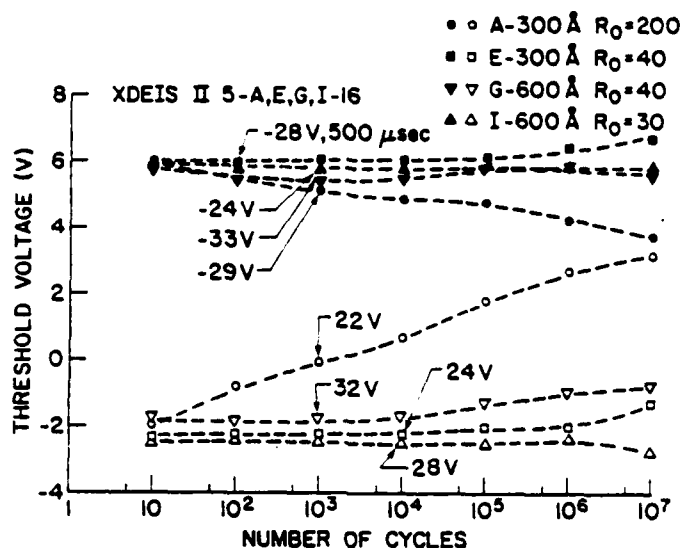
Fig. 18   Threshold voltage after write or erase operation as a function of the number of write/erase cycles on EAROM devices from the XDEIS II-5 series of wafers (see Fig. 1). Write/erase voltages as indicated in the figure were applied for 500 μsec. Wafer A has a 300 Å thick intervening layer of SiO_2 in the DEIS stack, while wafers E, G, and I have 300 Å or 600 Å thick intervening layers of OS-SiO_2 (containing $\simeq$ 3-6% excess atomic Si) in the MDEIS stack. The Si-rich SiO_2 injectors were 200 Å thick, and the gate oxide was 650 Å thick. These data show extended cyclability for MDEIS as compared to DEIS EAROMs due to the minimization of permanent trapped electronic charge.
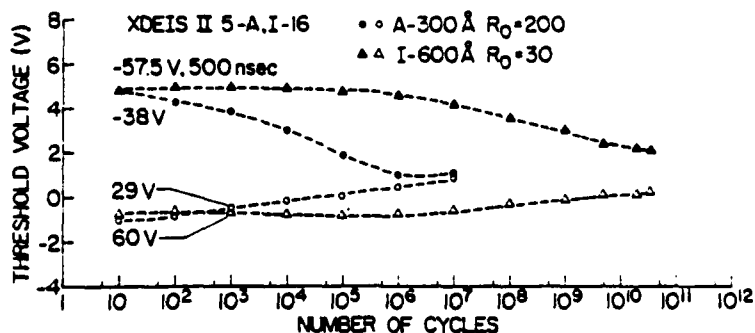
Fig. 19   Cycling characteristic for EAROM devices (XDEIS II 5-A, I-16) similar to that of Fig. 18 only for a write/erase pulse duration of 500 nsec and the voltages indicated in the figure. These data show that ~ 30% of the initial threshold voltage window is still left after ~ 4 x 10^{10} cycles for the MDEIS EAROM (I), and that the window has completely collapsed after ~ 10^6 cycles for the DEIS EAROM (A).

218

Fig. 20   Threshold voltage after write or erase operation as a function of the number of
write/erase cycles on MDEIS EAROMs (MDT-NVR 1-E-7), similar to that in Fig.
1 except for equal area floating and control gates, for various write/erase pulse
times from 5 msec to 5 μsec. Write/erase voltages are as indicated in the figure.
MDT-NVR 1-E-7 had a 100 Å thick gate oxide, 200 Å thick Si-rich SiO_2 injectors,
and a 600 Å thick intervening OS-SiO_2 in the MDEIS stack containing 5-6%
excess atomic Si (R_o=30).

Fig. 21   Threshold voltage after write or erase operation as a function of the number of
write/erase cycles on EAROMs (XDEIS II 5-E, F-16) with (E) and without (F)
200 Å thick Si-rich SiO_2 injectors. The OS-SiO_2 layer of both types of devices was
300 Å thick and contained 3-4% excess atomic Si (R_o=40). The gate oxide was
650 Å thick. Write/erase voltages as indicated in the figure were applied for 5
μsec.
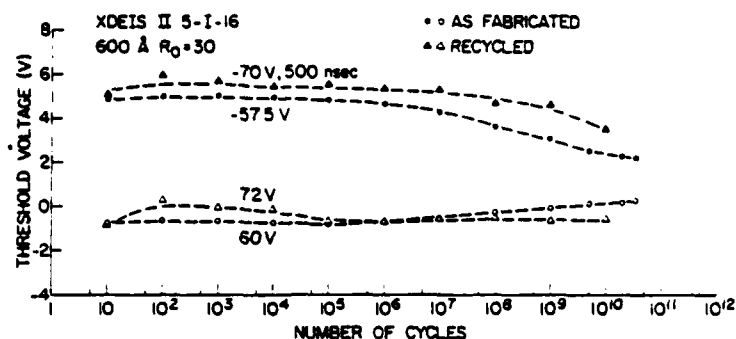
Fig. 22   Cycling characteristics for an MDEIS EAROM (XDEIS II 5-I-16 as described in the caption of Fig. 18) containing 5-6% excess atomic Si (R_o=30) in the 600 Å thick intervening OS-SiO_2 layer of the MDEIS stack which was cycled under conditions similar to those in Fig. 19 to $\sim 4 \times 10^{10}$ cycles and then recycled to 10^{10} cycles after the initial threshold voltage window was reestablished by increasing the write/erase voltage magnitudes as indicated in the figure.
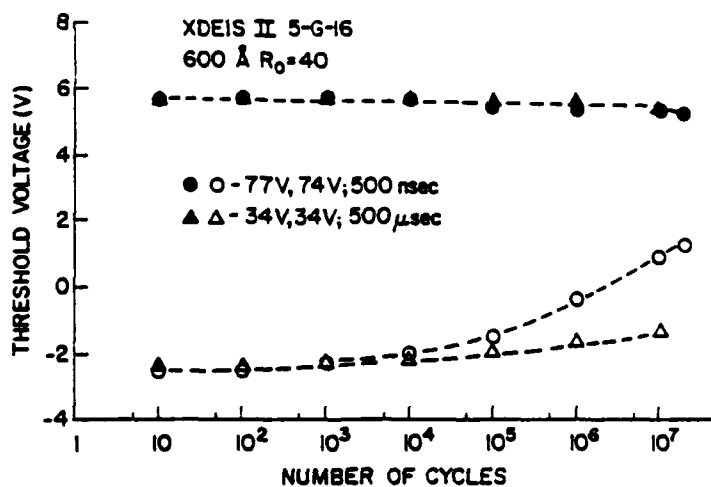
Fig. 23   Cycling characteristics on MDEIS EAROMs (XDEIS II 5-G-16 as described in the caption of Fig. 18) containing 3-4% excess atomic Si (R_o=40) in the 600 Å thick intervening OS-SiO_2 layer of the MDEIS stack which were cycled under different average electric field conditions (as indicated by the write/erase voltages and pulse times in this figure) from the same virgin as-fabricated threshold voltage window. This data show that for the same initial charge transferred between the floating and control gates, the device cycled under a very high field condition builds up permanent trapped electronic charge after fewer cycles.

220

DEIS II 2-CI-16
100Å R_0=200
• ○ -25 V, 20 V; 5 msec
■ □ -29 V, 24 V; 50 μsec
▲ △ -33 V, 26 V; 50 nsec
▼ ▽ -36.5 V, 27 V; 5 nsec
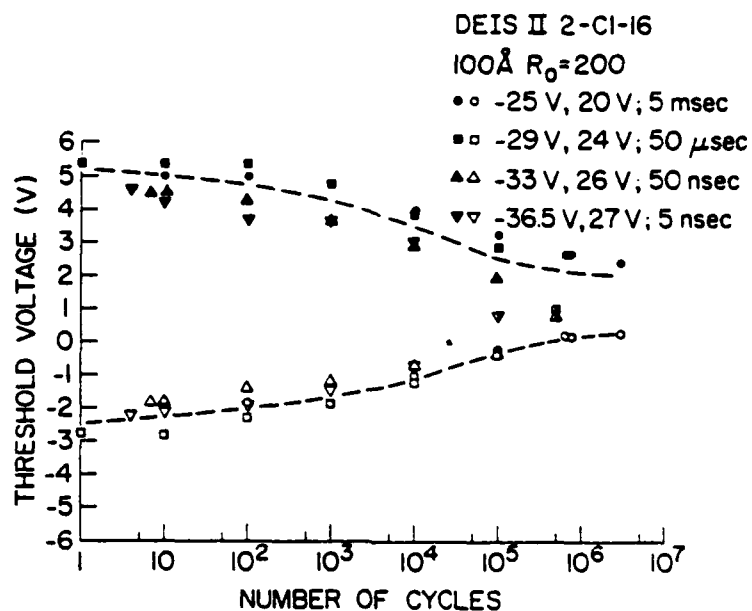
Fig. 24 Cycling characteristics on DEIS EAROMs (DEIS II 2-CI-16) containing a 100 Å thick intervening SiO_2 layer in the DEIS stack which were cycled under different average electric field conditions (as indicated by the write/erase voltages and pulse times in this figure) from the same virgin as-fabricated threshold voltage window. The Si-rich SiO_2 injectors were 150 Å thick, and the gate SiO_2 was 650 Å thick. These data show little effect of the average electric field on permanent trapped electronic charge build up in the SiO_2 layer of the DEIS stack.

XDEIS II 5-A,E,G,I-16
V_g=0 V, 25°C
○ A-300 Å R_0=200
□ E-300 Å R_0=40
▽ G-600 Å R_0=40
△ I-600 Å R_0=30
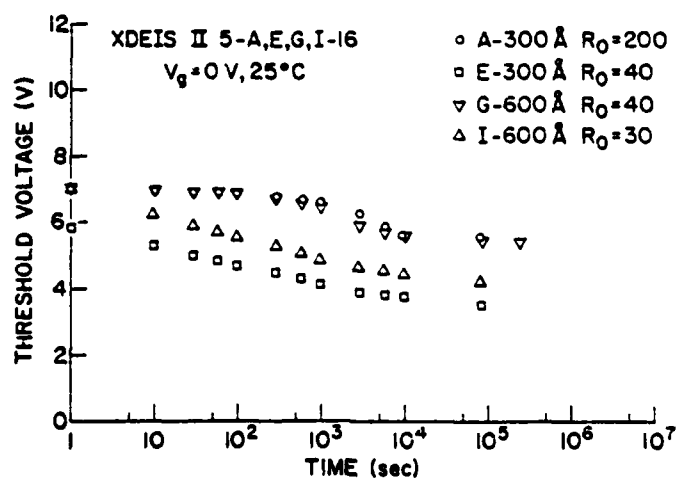
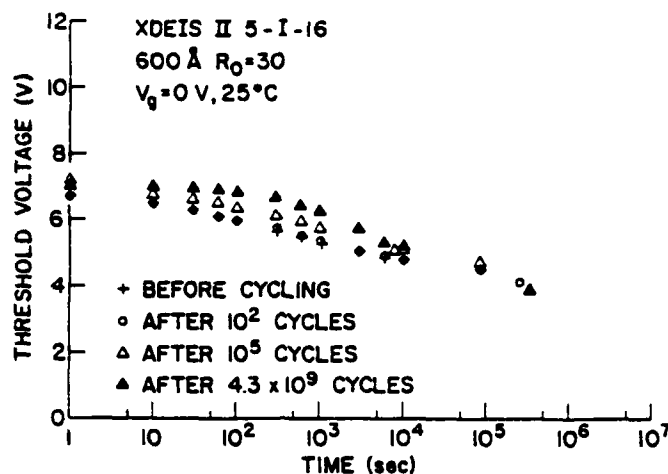**Fig. 25** Retention data (threshold voltage as a function of time) for a grounded control gate condition (0 V) at room temperature (25°C) on EAROM devices from the same series of wafers described in the caption of Fig. 18. The devices were charged from a virgin as-fabricated state to an initial threshold voltage of ≈ 6 to 7 V (written state). No correction for read disturb effects has been made.
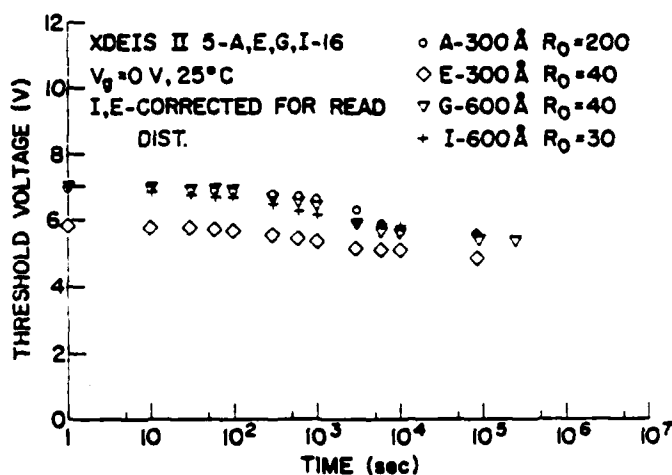
Fig. 26   Retention data of Fig. 25 corrected for read disturb.   This figure shows little increased charge loss with increasing Si content for the written state.
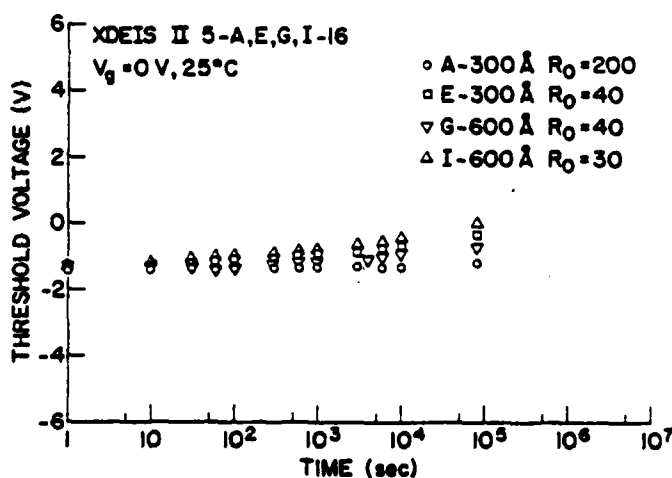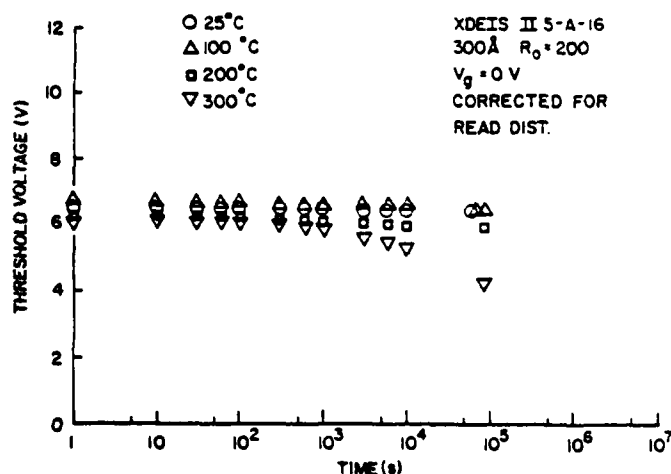

Fig. 27   Retention data (threshold voltage as a function of time) for a grounded control gate condition (0 V) at room temperature (25°C) on EAROM devices from the same series of wafers described in the caption of Fig. 18.   The devices were charged from a virgin as-fabricated state to an initial threshold voltage of $\approx$ -1.3 to -1.5 V (erased state).   Read disturb effects were negligible.   This figure shows little increased charge loss with increasing Si content for the erased stete.

Fig. 28   Retention data (threshold voltage as a function of time) for a grounded control
gate condition (0 V) at room temperature (25°C) on a MDEIS EAROM (XDEIS
II-5-I as described in the caption of Fig. 18) containing 5-6% excess atomic Si
(R_o=30) in the 600 Å thick intervening OS-SiO_2 layer of the MDEIS stack after
various amounts of cycling similar to that depicted in Fig. 19. This device was
charged from either a virgin as-fabricated state or a cycled state to an initial
threshold voltage of $\approx$ 7 V (written state). No correction for read disturb effects
has been made. This figure shows little increased charge loss with cycling for the
written state.

Fig. 29   Retention data (threshold voltage as a function of time) for a grounded control
gate condition (0 V) at various temperatures (25°C to 300°C) on DEIS EAROMs
(XDEIS II 5-A-16 as described in the caption of Fig. 18) containing a 300 Å thick
intervening SiO_2 layer in the DEIS stack. The devices were charged from a virgin
as-fabricated state to an initial threshold voltage of $\approx$ 6 to 7 V (written state).
The data were corrected for read disturb. These data show increased electronic
charge lost for the written state with increasing temperature.
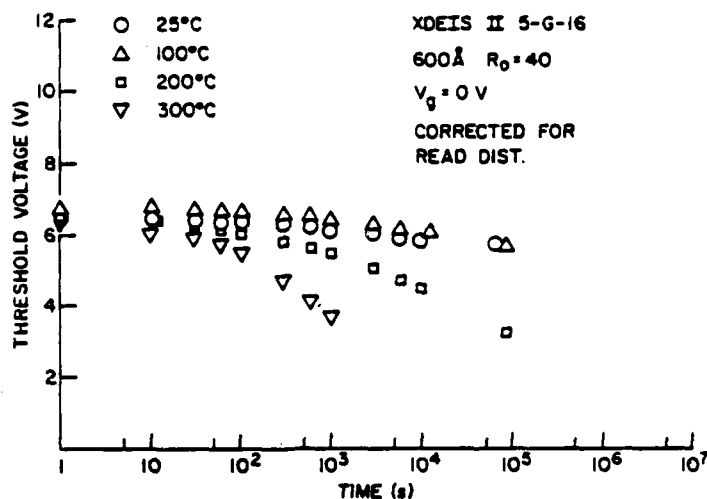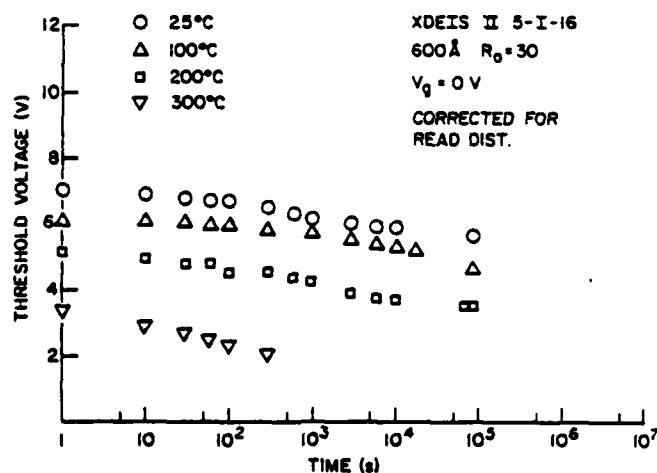
Fig. 30   Retention data at various temperatures under identical conditions to Fig. 29 on
MDEIS EAROMs (XDEIS II 5-G-16 as described in the caption of Fig 18) con-
taining 3-4% excess atomic Si (R_o=40) in the 600 Å thick intervening OS-SiO$_2$
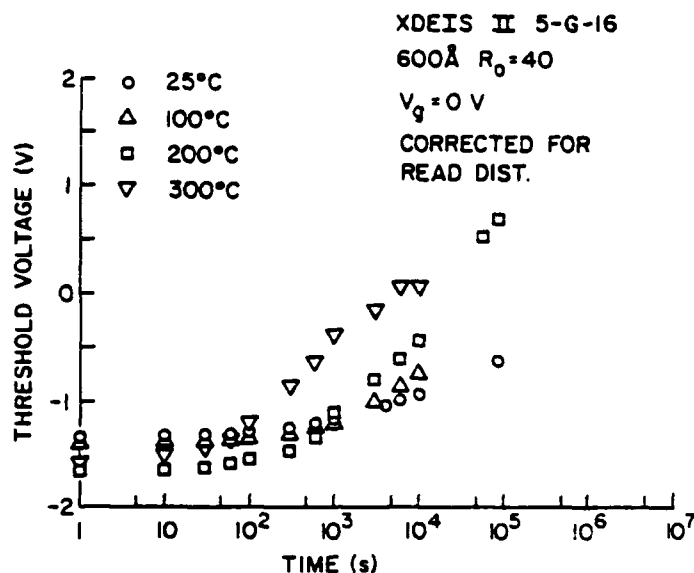layer of the MDEIS stack.

Fig. 31   Retention data at various temperatures under identical conditions to Fig. 29 on
MDEIS EAROMs (XDEIS II 5-I-16 as described in the caption of Fig. 18, con-
taining 5-6% excess atomic Si (R_o=30) in the 600 Å thick intervening OS-SiO$_2$
layer of the MDEIS stack.

Fig. 32   Retention data at various temperatures under similar conditions and on MDEIS
           EAROMs similar to those in Fig. 30 for devices which were charged from a virgin
           as-fabricated state to an initial threshold voltage of ≈ -1.3 to -1.5 V (erased state).
           These data show increased positive charge lost for the erased state with increasing
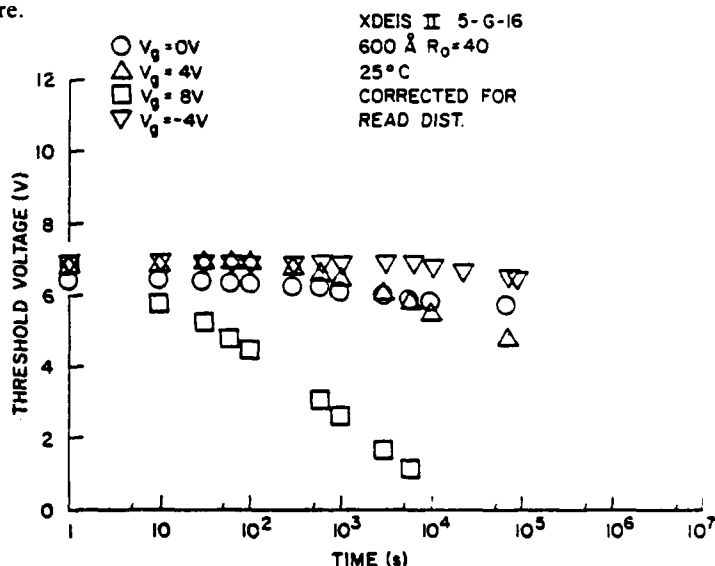   temperature.

Fig. 33   Retention data (threshold voltage as a function of time) for the control gate biased
           at various stressing voltages (-4 V to 8 V) at room temperature (25°C) on MDEIS
           EAROMs (XDEIS II 5-G-16 as described in the caption of Fig. 18) containing
           3-4% excess atomic Si (R_0=40) in the 600 Å thick intervening OS-SiO$_2$ layer of
           the MDEIS stack. The devices were charged from a virgin as-fabricated state to an
           initial threshold voltage of ≈ 6 to 7 V (written state). The data were corrected for
           read disturb. These data show increased electronic charge lost for the written state
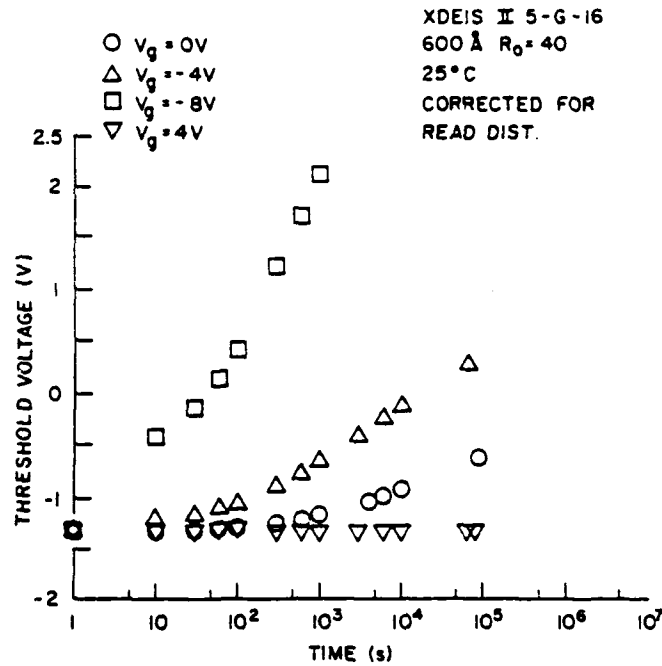   with increasing positive gate voltages.   225

Fig. 34   Retention data at various stressing voltages under similar conditions and on MDEIS
            EAROMs similar to those in Fig. 33 for devices which were charged from a virgin
            as-fabricated state to an initial threshold voltage of $\approx$ -1.3 V (erased state). These
            data show increased positive charge lost for the erased state with increasing nega-
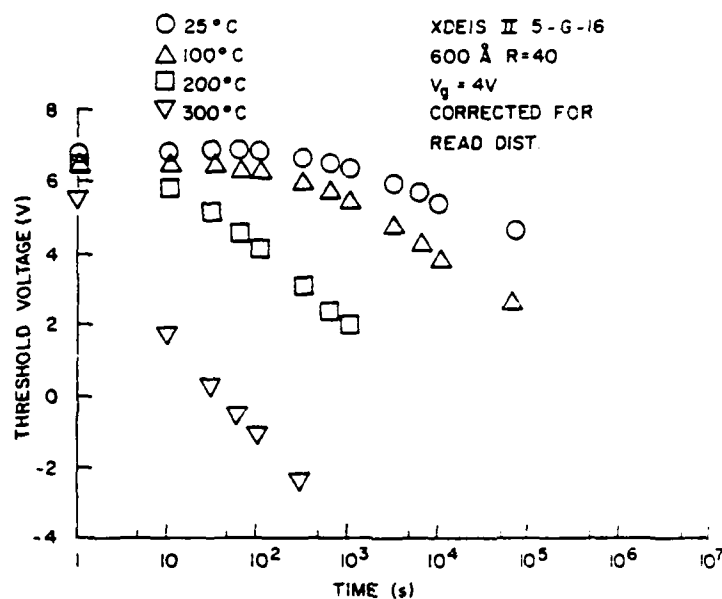   tive gate voltages.

53

Fig. 35   Retention data (threshold voltage as a function of time) for the control gate biased
at 4 V at various temperatures (25°C to 300°C) on MDEIS EAROMs (XDEIS II
5-G-16 as described in the caption of Fig. 18) containing 3-4% excess atomic Si
(R_o=40) in the 600 Å thick intervening OS-SiO_2 layer of the MDEIS stack. The
devices were charged from a virgin as-fabricated state to an initial threshold voltage
of ≈ 6 to 7 V (written state). These data were corrected for read disturb. These
data show increased electronic charge lost for the written state with increasing
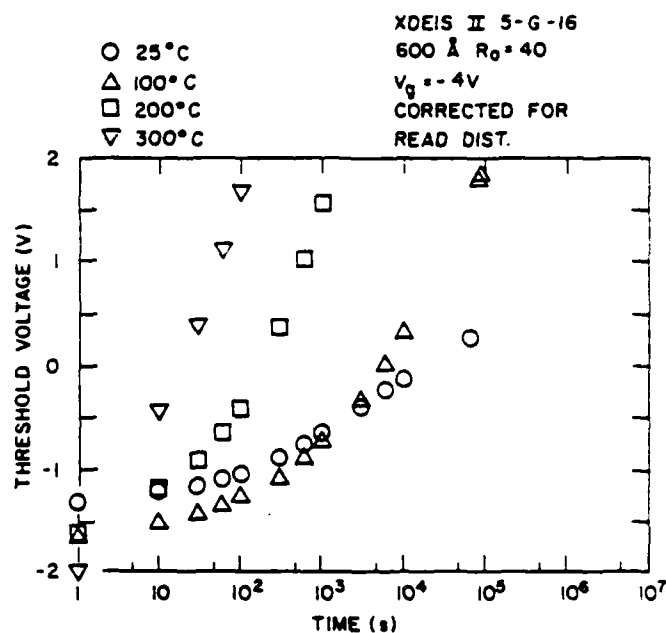temperature with a 4 V stressing voltage applied to the gate electrode.

227

Fig. 36   Retention data for the control gate biased at -4 V at various temperatures on
MDEIS EAROMs similar to those in Fig. 35 for devices which were charged from a
virgin as-fabricated state to an initial threshold voltage of $\approx$ -1.3 to -1.5 V (erased
state).   These data show increased positive charge lost for the erased state with
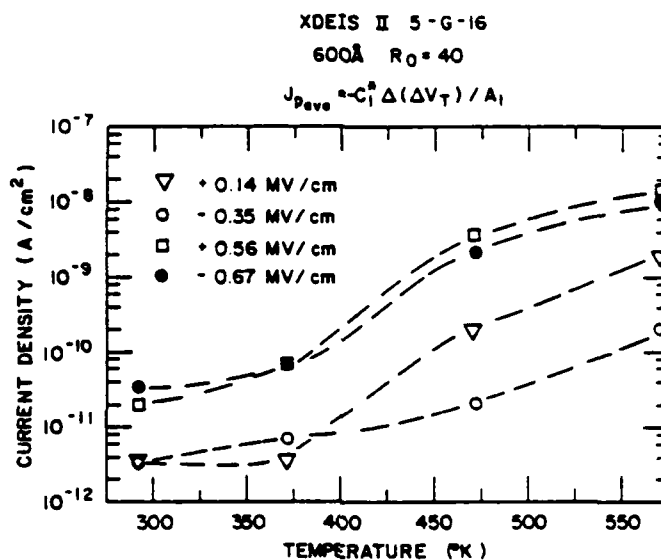increasing temperature with a -4 V stressing voltage applied to the gate electrode.

Fig. 37   Magnitude of the average current density as a function of temperature for various low electric field conditions deduced using the retention characteristics of MDEIS EAROMs (XDEIS II 5-G-16 as described in the caption of Fig. 18) containing 3-4% excess atomic Si (R_o=40) in the 600 Å thick intervening OS-SiO$_2$ layer of the MDEIS stack. Calculations were performed as described in the text (see Section III-B-3) using the data from Fig. 30 (∇), Fig. 32 (O), Fig. 35 ($\square$), and Fig. 36 ($\bullet$).

MDT - NVR I - E - 7
600 Å $R_0 = 30$;
$V_T(0) = 2.6$ V; WRITE V_g^-
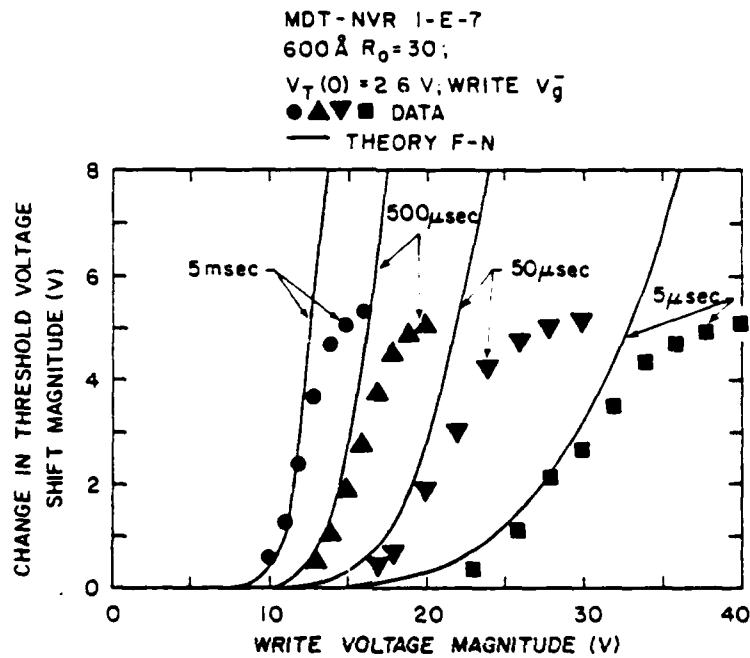● ▲ ▼ ■ DATA
—— THEORY F-N

Fig. 38   Comparison of model calculations described in Section IV with experimental data for the change in the threshold voltage shift magnitude as a function of write voltage magnitude on a MDEIS EAROM (MDT-NVR 1-E-7 as described in the caption of Fig. 20) containing 5-6% excess atomic Si ($R_0 = 30$) in the 600 Å thick intervening OS-SiO$_2$ layer of the MDEIS stack. The write pulse was applied for various times (5 msec to 5 µsec) from an initial $V_T(0) = 2.6$ V condition which corresponds to the floating gate in an uncharged state.

Ellipsometry Measurements of Polycrystalline Silicon Films

E. A. IRENE AND D. W. DONG

IBM Corporation, Thomas J. Watson Research Center, Yorktown Heights, New York 10598
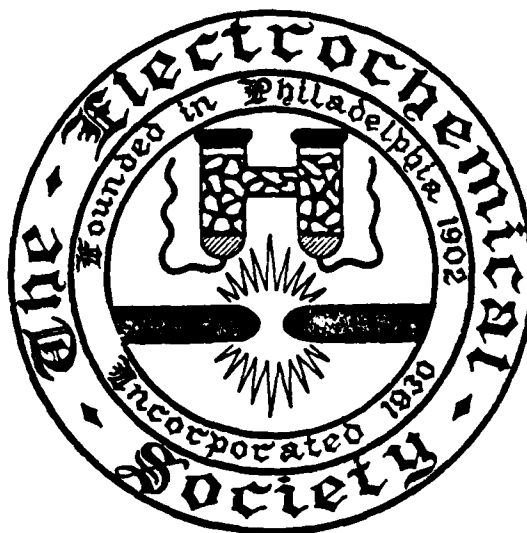
Ellipsometry Measurements of Polycrystalline Silicon Films

E. A. Irene* and D. W. Dong*

IBM Corporation, Thomas J. Watson Research Center, Yorktown Heights, New York 10598

ABSTRACT

This study shows that visible light ellipsometry can be used to measure the thickness of polycrystalline silicon films that have been deposited on oxidized Si slices. The measurement hinges on obtaining a value for the complex index of refraction, $\tilde{N}$, for the polycrystalline silicon films. An empirical method has been found which gives usable values for $\tilde{N}$. The limits of application of the ellipsometric technique are shown. Also, it is shown that discontinuous polycrystalline silicon films can be detected using ellipsometry. Finally, although theoretically possible, the accurate measurement of oxide grown on polycrystalline silicon surfaces is not practical.

The use of ellipsometry to measure the film thickness of polycrystalline silicon (poly-Si) films and silicon dioxide films grown on polycrystalline silicon (poly-OX) has been reported (1, 2). Dell'Oca (3) has listed three major drawbacks with the use of ellipsometry for poly-Si film measurements. First, the ellipsometric period for poly-Si is small (about 80 nm for 632.8 nm light) and therefore the poly-Si thickness must be known to $\pm$ 80 nm for 632.8 nm light and even closer for shorter wavelength light. Second, poly-Si films are usually deposited on SiO_2 films and therefore the SiO_2 film thickness must be accurately known. Third, poly-Si surface roughness may alter the polarization state of the reflected light and thereby render the system too complex for straightforward analysis. In the present study we have found that while all of Dell'Oca's points have some merit, successful measurements of poly-Si film thickness can be made using ellipsometry. It will be demonstrated that the first essential task is to measure the complex index of refraction, $\tilde{N}$, for poly-Si. An empirical method will be shown to yield reliable values for $\tilde{N}$. Next, the limits and applications of the measurement of poly-Si films will be presented in terms of annealing effects, thickness effects, and oxidation behavior.

With reference to Dell'Oca's first point above, ellipsometric periods for poly-Si (using $\tilde{N}$ obtained later) do not repeat identically in Δ, ψ space. Figure 1 shows a plot of Δ and ψ for poly-Si film thicknesses up to about 600 nm. It is seen that the ability to discriminate between ellipsometric periods depends on the ability to measure Δ and ψ accurately and the knowledge of $\tilde{N}$

for poly-Si. Since commercially available high quality ellipsometers can routinely measure Δ and ψ to better than 0.1°, instrumental accuracy is not considered in this study. The real problem is to have an accurate reproducible measurement of $\tilde{N}$ for poly-Si. The pres-
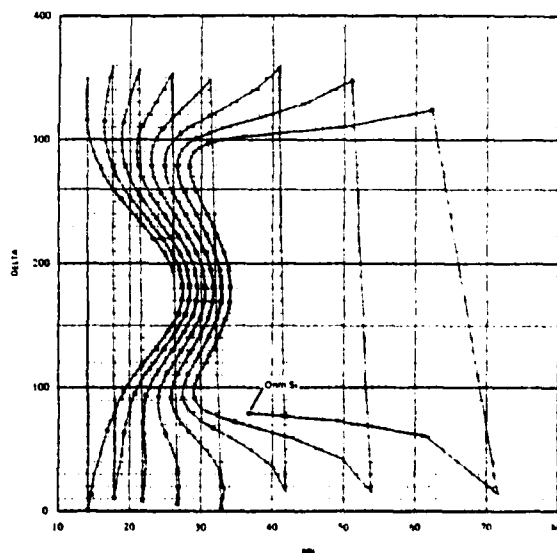

Fig. 1. Theoretical plot of delta and psi as a function of poly-Si thickness (large dots are 5 nm poly-Si thickness) on 90 nm SiO_2 on a single crystal Si substrate. $\tilde{N}$ for poly-Si is taken to be 4.06 (1-0.012i) for 632.8 nm light and an angle of incidence, ϕ, cf 70°.

* Electrochemical Society Active Member.
Key words: visible light ellipsometry, film thickness, discontinuous films, oxidation.

232

ent study is directed toward this measurement. An empirical procedure is employed which is then tested and the limitations discussed.

A single ellipsometric measurement yields values for the two parameters Δ and ψ. Therefore, two other parameters of the measured system are accessible via calculation. For the typical case of a poly-Si film on an SiO_2 film all on a single crystal Si substrate, the measured Δ, ψ value can be used to obtain two desired parameters providing all the others are known. For the Si substrate surface and each film, there exists a complex refractive index $\widetilde{N}$ as

$$\widetilde{N} = N(1 - i\kappa)$$

where N is the refractive index and κ is called the absorption index. Additionally, for each film there is the optical path length or film thickness. The values of $\widetilde{N}$ for single crystal silicon and SiO_2 films are well known.

Therefore, if $\widetilde{N}$ for the poly-Si film is known, then both SiO_2 and poly-Si film thickness can be obtained from a single ellipsometric measurement and Dell'Oca's second objection can be overcome. In practice, it will be shown that a usable value of $\widetilde{N}$ for poly-Si can be obtained, but the error associated with $\widetilde{N}$ will be larger than the errors of $\widetilde{N}$ for the substrate or the SiO_2 film and, therefore, an independent measurement of the thickness of the underlying SiO_2 film is desirable so as to over specify the system. It would also be useful to use ellipsometry to measure the film thickness of SiO_2 films grown via oxidation of poly-Si (herein these oxide films are referred to as poly-OX). Such ellipsometric measurements have been reported in the literature (1), but the specific procedures and limitations of these measurements were not discussed. From measurements of the SiO_2 underlying poly-Si, $\widetilde{N}$ for poly-Si, and the poly-Si thickness, the poly-OX thickness can in principle be measured. However, the question arises of the effect of surface roughness on ellipsometric measurements. It is known that the poly-Si surface is rough due to grain growth and this surface becomes considerably rougher upon oxidation (4, 5). In the present study, poly-OX thickness measurements are compared to a theoretical calculation. It is shown that deviations from theory can be accounted for by both errors in poly-Si thickness and surface roughness.

Experimental Procedures

Silicon films were deposited by chemical vapor deposition (CVD) at 650°C onto oxidized single crystal silicon slices. At that temperature, we have found that the silicon films are composed of small crystals in an amorphous silicon medium (4). Later annealing treatments at 1000°C for a few minutes converts the silicon to polycrystalline silicon. The oxide grown on the single crystal silicon was grown at 1000°C in dry O_2. The single crystal silicon slices were commercially available, (100) oriented, chem-mechanically polished 2 Ω-cm, device quality silicon.

The ellipsometric measurements were made using either a Rudolf Research Type 43603-200E thin film ellipsometer using 546.1 nm light and equipped with a quarter-wave plate or an automated ellipsometer (6) using 632.8 nm light. Both of these instruments have a capability of 0.01° accuracy in polarizer and analyzer settings. In practice, we found that the overall precision based on repeated thickness measurements using the manual ellipsometer was always better than 2% and about 3% for the automated ellipsometer. Theoretical ellipsometric calculations as well as calculations based on experimental data were made using McCrackin's program (7).

The relevant ellipsometric relationships are obtained by first considering the polarization states of single waves reflected from a sample surface that is covered with one or several films and then the total reflected amplitude, R, of many waves (a detailed treatment can be found in Ref. (8)).

For reflected waves parallel (p) and perpendicular (s) to the plane of incidence, the complex Fresnel coefficients can be written as

$$r_p = |r_p| e^{i\delta_p} \qquad [1]$$

and

$$r_s = |r_s| e^{i\delta_s} \qquad [2]$$

where $|r|$ is the amplitude ratio and δ is the phase shift upon reflection. For many waves the sum of all r_p and r_s components emanating from a number of surfaces yields R_p and R_s of the same form. Reflection ellipsometry endeavors to measure the ratio

$$\rho = R_p/R_s \qquad [3]$$

This ratio is also conveniently written as

$$\rho = \tan \psi \, e^{i\Delta} \qquad [4]$$

where

$$\tan \psi = |R_p|/|R_s| \qquad [5]$$

and

$$\Delta = \Delta_p - \Delta_s \qquad [6]$$

Thus, ψ and Δ are related to changes in the amplitude and phase of the incident waves upon reflection. ψ and Δ are measurables in ellipsometry and they are obtained from polarizer, compensator, and analyzer instrumental settings in a straightforward manner. Of particular relevance is the relationship of the ellipsometric measurables, ψ and Δ, to the desired optical properties of the film-substrate combinations. These desired properties are the refractive index $\widetilde{N}$ and thickness d for each film and the index for the substrate. This relationship is obtained from the explicit form for the reflection coefficients [see for example Ref. (8)]. The result is summarized by the relationship

$$\tan \psi \, e^{i\Delta} = \rho \, (\widetilde{N}_i, d_i, \phi, \lambda) \qquad [7]$$

where $\widetilde{N}_i$ and d_i are the various complex indexes and film thicknesses for each film, ϕ is the angle of incidence, and λ is the wavelength of the incident light. Equation [7] is inverted using McCrackin's program (7). It should be observed that with only two actual measurables, only two optical properties are directly obtainable. Therefore, for the case of multiple films on a substrate one measurement (at a constant ϕ and λ) will not yield all the d and $\widetilde{N}$ values. It should also be noticed that for absorbing films, κ is nonzero and $\widetilde{N}$ for such a film has two parts. This means that a single measurement for an absorbing film (at constant ϕ and λ) cannot yield all the optical properties. N, κ, and d, for the film. However, empirical estimates to obtain additional properties from a single measurement can be made and one such estimation procedure is described below.

In order to confirm a number of the ellipsometric measurements made during this study, two other conventional film thickness measurement techniques were used: the normal incidence interference technique and the mechanical step height technique.

Experimental Results

Measurement of $\widetilde{N}$.—Numerical values of $\widetilde{N}$ for poly-Si films were obtained by an empirical procedure. For this measurement, silcon films of various approximately known thicknesses were deposited on SiO_2 films with previously ellipsometrically measured film thick-

ness all on Si single crystal substrates. These samples were annealed at 1000°C for 30 min in N_2 to insure crystallization of the silicon films, and any surface oxide grown on the poly-Si films as a result of annealing was removed by a few seconds dip in HF. The justification for the anneal to insure crystallization is treated separately below. The experimentally measured Δ and ψ values for the poly-Si films were then plotted along with theoretical Δ, ψ curves for the layered structure under study as a function of poly-Si thicknesses. Figure 2, a through e, shows the essential features of this type of curve fitting where the theoretical curves were calculated for various values of $\tilde{N}$. For 632.8 nm light, a best fit was seen at $N = 4.06$, $\kappa = 0.012$ (Fig. 2b). It is seen that for too low a value for N, the data does not lie on any theoretical curve, but for too high N, the data falls on or near second and third period theoretical curves. Independent film thickness measurements (optical interference and mechanical step height) established that the film thicknesses were all less than 160 nm. A similar procedure for measurements taken with 546.1 nm light yielded values of $N = 4.30$ and $\kappa = 0.024$. As a further check whether this empirical procedure leads to usable values for $\tilde{N}$, a set of poly-Si samples on oxide were grown from 70 to 450 nm thick. The measured Δ, ψ values are plotted along with the theoretical curve in Fig. 3. All the samples up to 200 nm are close to the theoretical curve thereby confirming that usable values for $\tilde{N}$ are obtained.

For the measured value of $\tilde{N}$ for poly-Si, it is seen that adjacent ellipsometric periods are about 2° apart in ψ (see Fig. 1 or 2). Therefore, better than 1° accuracy in ψ is needed to reasonably locate the period. The instruments used in the present study, when properly aligned and calibrated, can measure Δ and ψ to better than 0.1°. Therefore, the location of the proper period will in large part be a function of how good was the selection of $\tilde{N}$. To demonstrate the effects of possible errors associated with the empirical curve fitting, Table I was calculated. Table I shows the effect of differences in $\tilde{N}$ values as $\pm$ 10% in N and $\pm$ 10% and $\pm$ 20% in κ ($\pm$ 100% for N = 4.06) on resultant poly-Si film thicknesses. It is seen that errors in κ of 20% hardly affect the film thickness, d, and 10% errors in N affect d by about 10%. It is clear from Fig. 2 that errors of less than 10% in N and 100% in κ are made. From Table I, errors in the film thickness for these deviations are about 10% or less. Therefore, we can conclude that this procedure is adequate to better than 10% in terms of the resultant poly-Si film thickness, although κ values may be in greater error.

Annealing behavior of poly-Si films.—The Δ, ψ measurements made in this study on the silicon films de-

posited at 650°C strongly suggest that the films possess both different and irreproducible values of $\tilde{N}$ as compared with samples annealed above 700°C. An illustration of this annealing effect is shown in Fig. 4 where it is seen that samples grown in the same CVD run and therefore having about the same film thickness (65-70 nm) have Δ, ψ values not only substantially different for each sample but also to the left (lower ψ) of a theoretical curve calculated using $\tilde{N}$ derived previously. However, upon annealing these samples at 1000°C in N_2 for about 15 min the samples fall in a small grouping very near the theoretical curve. As was demonstrated in a previous study (4), the silicon films deposited at 650°C are composed of rather small crystallites embedded in an amorphous silicon matrix. Since it is known that amorphous silicon has both larger N and κ values than crystalline silicon, it would be useful to determine if the lower values of Δ and ψ and for some samples somewhat higher Δ values for the as-deposited material are a result of the larger N and κ values. The results of calculations in Table II show that as N increases, Δ decreases substantially while ψ increases slightly. For an increase in κ, both Δ and ψ decrease. Therefore, it is likely that the smaller values of ψ for the as-deposited silicon is due to the higher κ values for this nearly amorphous material. The scatter in Δ for unannealed samples is not explainable at this time. For some CVD runs, we have noticed Δ values larger, sometimes the same, and sometimes smaller than the annealed values. There may be a relationship between the Δ values and the exact preparation conditions; however, this has not been adequately investigated.

Discontinuous poly-Si films.—Up to this point, the present study has shown that a reasonable measurement of $\tilde{N}$ could be made on poly-Si films thereby enabling the measurement of film thickness for thin poly-Si films. In order to determine the minimum thickness that could be detected, very thin poly-Si films were deposited on already prepared SiO_2 films on Si. Thickness estimations were obtained by extrapolation based on the CVD times needed to grow thicker films. Δ, ψ data for films estimated to be as thin as 2 nm and thicker are shown along with an SiO_2 on Si curve and a poly-Si curve in Fig. 5. The SiO_2 and poly-Si curves intersect at the SiO_2 thickness which exists under the deposited poly-Si but the 10 sec run would correspond to 70.5 nm of SiO_2 rather than the 57.5 nm of starting SiO_2. From points corresponding to 20, 30, and 60 sec CVD runs, it is seen that the data scatter near the theoretical curve. For the 180 sec and longer runs, the data fall on the theoretical curve. Therefore, it appears likely that the coverage of the film between 10 and 180 sec is less than 100%. To confirm this result, TEM was performed. Figure 6 for a 2 min run shows that this film has less than 100% coverage, although it was found to be electrically continuous as is evidenced by the connectivity of the poly-Si islands.

Oxidation of poly-Si films.—With a knowledge of the thickness of a poly-Si film on SiO_2, it is also possible to measure the growth of an oxide film on poly-Si using

Table I. The effects of ±10% variation in N and ±20 and ±100% variation in κ on poly-Si film thickness

N	κ	Calculated film thickness (nm)
3.65	0.0108	56.2
	0.0120	56.2
	0.0132	56.2
4.47	0.108	45.0
	0.120	44.9
	0.132	44.9
4.06	0.0060	49.9
	0.0120	49.8
	0.0240	49.6

Conditions: 30 nm poly-Si on 50 nm SiO_2 on Si, λ = 546.1 nm, ϕ = 70°, Δ = 160.722, ψ = 30.922.

Table II. The effect of changes in N and κ on Δ and ψ values

N	κ	Δ	ψ
4.0	0	163.3	30.8
	0.1	159.1	24.5
	0.2	153.1	21.0
4.1	0	159.0	31.0
4.2	0	153.9	31.2

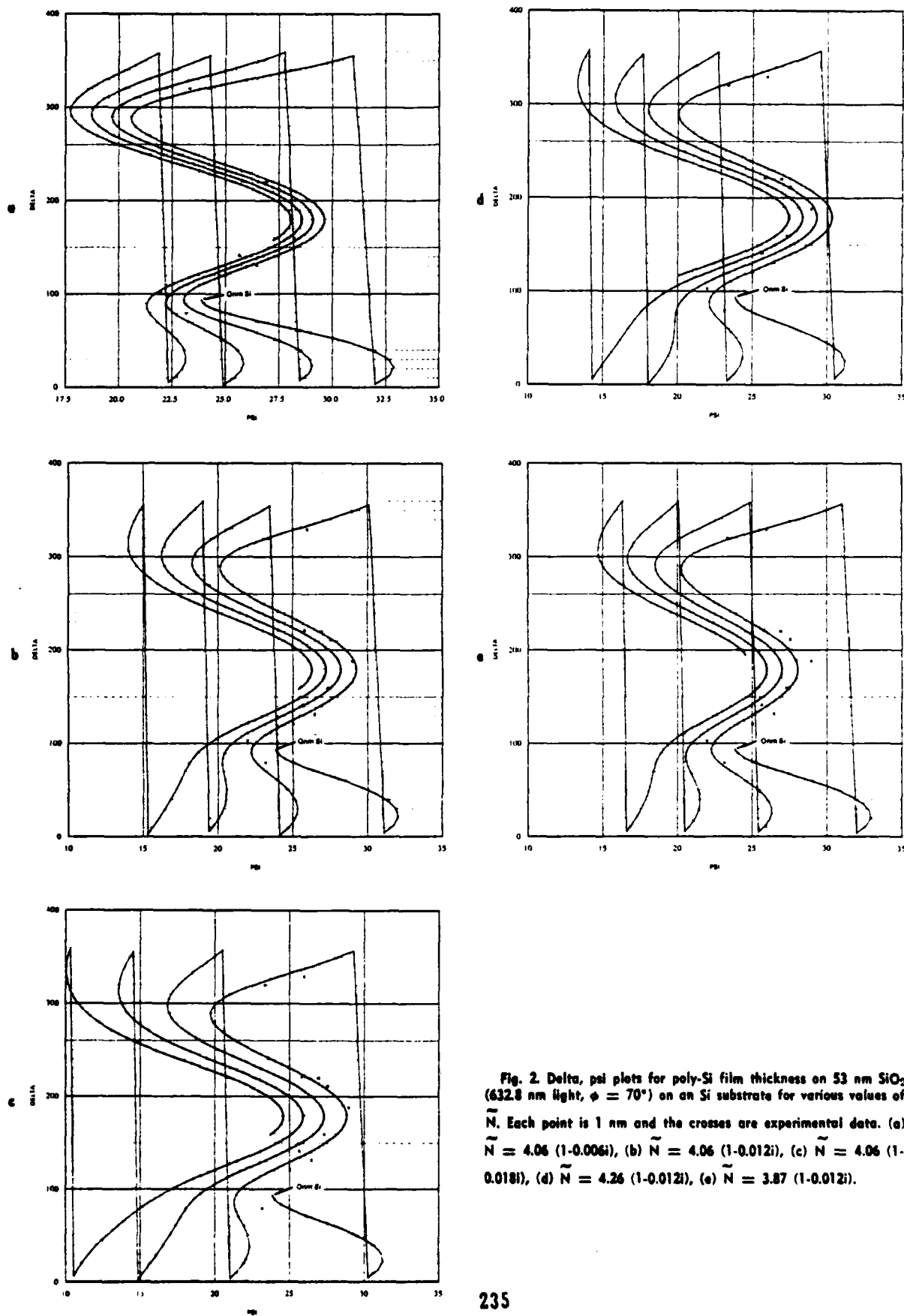Conditions: 50 nm poly-Si on 50 nm SiO_2 on Si, λ = 546.1 nm, ϕ = 70°.

Fig. 2. Delta, psi plots for poly-Si film thickness on 53 nm SiO$_2$ (632.8 nm light, $\phi = 70°$) on an Si substrate for various values of $\widetilde{N}$. Each point is 1 nm and the crosses are experimental data. (a) $\widetilde{N} = 4.06$ (1-0.006i), (b) $\widetilde{N} = 4.06$ (1-0.012i), (c) $\widetilde{N} = 4.06$ (1-0.018i), (d) $\widetilde{N} = 4.26$ (1-0.012i), (e) $\widetilde{N} = 3.87$ (1-0.012i).
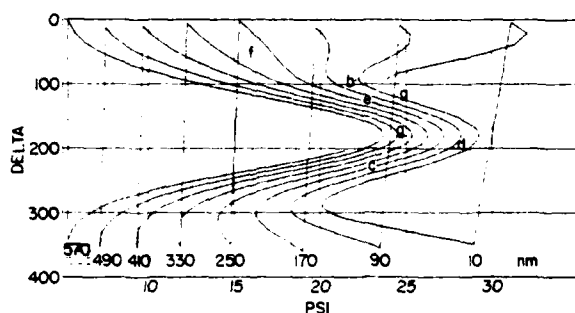
235

Fig. 3. Theoretical delta, psi plots for poly-Si film thicknesses on 53 nm SiO₂ (632.8 nm light, φ = 70°) for N̄ = 4.06 (1-0.012i). Each point is 1 nm on the theoretical curve. The letters indicate experimental data: a = 72 nm, b = 85 nm, c = 142 nm, d = 204 nm, e = 310 nm, f = 324 nm, and g = 455 nm.
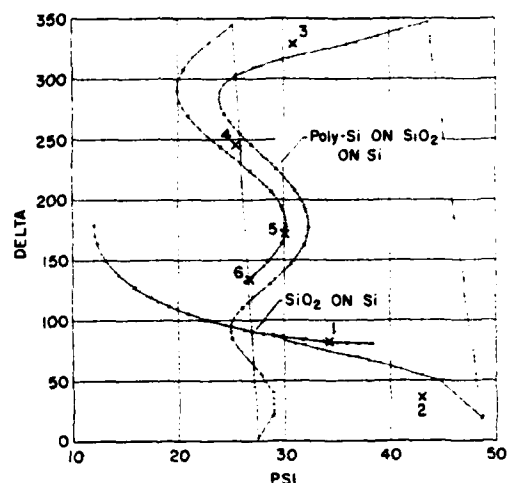

Fig. 5. Theoretical delta, psi plots of poly-Si on 57.5 nm SiO₂ on Si with N̄ = 1.465 (1-0i) for 546.1 nm light. The experimental points (crosses) are for sequentially longer poly-Si deposition times: 1 = 10 sec, 2 = 20 sec, 3 = 30 sec, 4 = 60 sec, 5 = 180 sec, and 6 = 300 sec. All samples were annealed at 1000°C, 10 min, N₂ after deposition and briefly dipped in buffered HF to remove surface oxide.

Fig. 4. Theoretical delta, psi plots for 106 nm SiO₂ on Si (solid curve) with data points for unannealed poly-Si films (open circles) and annealed (1000°C, 30 min, N₂) poly-Si (crosses). The two sets of data appear for unannealed samples due to a degeneracy in the computation.

Fig. 6. TEM micrograph of 120 sec deposited poly-Si film on SiO₂

ellipsometry. In order to accomplish this measurement, the following algorithm was employed. First, the oxide that will underlay the poly-Si is thermally grown on a single crystal Si substrate and measured by ellipsometry. Then the poly-Si is deposited on this oxide, annealed, HF dipped to remove any surface oxide, and *the poly-Si thickness is measured as described above.* The sample is then placed in the oxidation furnace of an automated ellipsometer [previously described in Ref. (6)] and brought to oxidation temperature in N₂. After the sample is thermally equilibrated, Δ, ψ data are taken as the initial value of the oxidation run. The ambient is switched to oxygen and Δ, ψ are taken at regular time intervals. After a desired oxidation time, the sample is returned to room temperature in N₂ and a final Δ, ψ value is measured. The poly-OX is removed in HF and the remaining poly-Si film thickness is measured. From this measured poly-Si value and the final Δ, ψ value taken before the poly-OX was chemically removed, the final poly-OX thickness is calculated. From the starting and final poly-Si film thicknesses and the poly-OX thickness, the molar volume ratio for poly-OX is calculated to be for Si/SiO₂ = 0.44. From published

density values for Si and SiO₂ (9), the molar volume ratio is calculated to be 0.456. Considering about 3% error in the poly-Si and poly-OX measurements and variability in density values, the measured and calculated ratios are in excellent agreement. Now for the first measured Δ, ψ value after O₂ was turned on, we used the initial value for the poly-Si thickness and the measured Δ, ψ value to calculate the first value for poly-OX thickness. From this value for poly-OX and

the molar volume ratio, the next value for poly-Si is calculated and is used along with the next Δ, ψ value to calculate the next poly-OX, etc. This process continues until all the measured Δ, ψ values yield poly-Si and poly-OX thickness values. This method is easily checked since the final calculated values must compare with the final measured values where the final poly-Si value was measured independently. Agreement was always found to be better than 10%.

Figure 7 shows a plot of the oxidation data obtained as a result of the above calculational procedure. The increase in poly-OX and the decrease of poly-Si film thicknesses is seen. The procedure used to calculate these thickness-time relationships is likely to be quite accurate at the end and beginning of the oxidation, since individual film measurements are made. However, intermediate measurements depend on the measured Δ, ψ (which is in general quite precise) and the adherence of the progressive oxidation to a theoretical curve. Below is a discussion of these intermediate measurements.

Figure 8a shows the changes in Δ and ψ as poly-Si is being oxidized. It is seen that at the beginning of the oxidation the experimental data lie close to the theoretical curve. As oxidation proceeds, however, the data deviate from the curve and when oxidation of the poly-Si is nearly complete, the experimental points approach the SiO2 theoretical curve. Figure 8b shows four theoretical poly-Si oxidation curves plus experimental data. One poly-Si and oxide curve is calculated for an initial poly-Si thickness of 44.5 nm (which corresponds to the real sample) and the others for deviations of the initial poly-Si in 2.0 nm increments of starting poly-Si. It is seen from this figure that a 2.0 nm difference in starting poly-Si thickness translates into a rather large change in the theoretical curves for the oxidation of poly-Si. The theoretical curve for an initial poly-Si thickness of 44.5 nm fits the experimental data best prior to and after complete oxidation. During oxidation, the 46.5 nm curve fits closely. The conclusion is that the error limits of about ± 10% in poly-Si thickness translates to quite a large error in the oxide in poly-Si measurements. Since the experimental data seem to be somewhat asymmetrical with respect to the theoretical poly-Si oxidation curves, other factors such as surface roughness caused by intergranular oxidation (4, 5) are likely to be affecting the measurement. Whatever the cause, we believe that since the measurement of poly-Si is no better than ± 10%, and since the measurement of poly-OX depends on a precise poly-Si thickness, the measurement of poly-OX appears to be in error by substantially more than 10%.
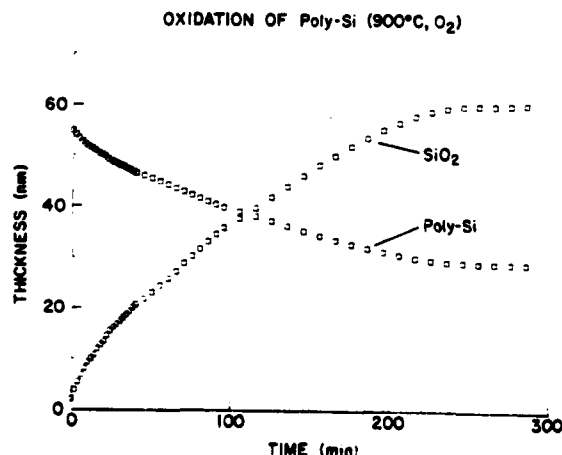

Fig. 8. Theoretical delta, psi plots of the oxidation of poly-Si and for SiO2. (a) Undoped poly-Si, 44 nm thick on 140 nm SiO2. The crosses are experimental data. (b) Doped poly-Si, 45 nm thick on 90 nm SiO2 (crosses). The four poly-Si theoretical curves are for 2 nm differences in starting poly-Si thicknesses: 40.5 (triangles), 42.5 (circles), 44.5 (squares), and 46.5 (diamonds).

Conclusions

Visible light ellipsometry can be used to measure the thicknesses of poly-Si films on oxidized silicon slices.

The measurement hinges on the knowledge of $\tilde{N}$ for the silicon film. For annealed poly-Si films, $\tilde{N}$ is obtainable by an empirical procedure. For large poly-Si thicknesses, errors in $\tilde{N}$ will prevent accurate measurement, but for films less than several thousand angstroms, the ellipsometric technique is usable. The ellipsometric technique can be used to detect less than 100% film coverage without resorting to destructive techniques such as TEM. In principle, it should be possible to measure the oxidation of poly-Si since the poly-Si thickness can be measured. However, the combined effects of errors in the poly-Si film thickness and intergranular oxidation, which roughens the poly-Si surface, obfuscates the measurement.

OXIDATION OF Poly-Si (900°C, O2)

Fig. 7. Poly-OX and poly-Si film thickness vs. time of oxidation at 900°C in pure O2. The thicknesses are calculated from experimental data and the procedure outlined in the text.
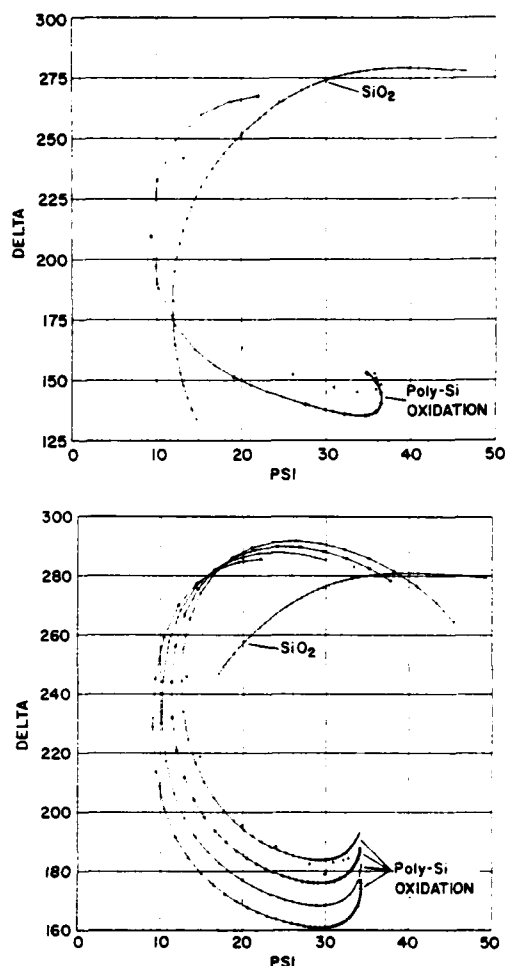
Acknowledgments

REFERENCES

1. T. I. Kamins and E. L. McKenna, *Metal Trans.*, **2**, 2292 (1971).
2. T. I. Kamins and R. R. Cass, *Thin Solid Films*, **16**, 147 (1973).
3. C. J. Dell'Oca, *This Journal*, **119**, 108 (1972).
4. E. A. Irene, E. Tierney, and D. W. Dong, ibid., **127**, 705 (1980).
5. E. A. Irene and E. Tierney, "Microstructural Science," Vol. 8, p. 91, Elsevier North Holland Inc. (1980).
6. Y. J. van der Meulen and N. C. Hien, *J. Opt. Soc. Am.*, **64**, 804 (1974).
7. F. L. McCrackin, Natl. Bur. Std., Technical Note 479, U.S. Govt. Printing Office, Washington, DC (1969).
8. R. M. A. Azzam and N. M. Bashara, "Ellipsometry and Polarized Light," North Holland Co., Amsterdam (1977).
9. A. S. Grove "Physics and Technology of Semiconductor Devices," John Wiley & Sons, Inc., New York (1967).

Silicon Oxidation Studies: Measurement of the Diffusion of Oxidant in SiO₂ Films

E. A. IRENE

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

Silicon Oxidation Studies: Measurement of the Diffusion of Oxidant in SiO₂ Films

E. A. Irene*

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

ABSTRACT

A method for the measurement of the diffusion of oxidant through a growing SiO₂ film is presented. The procedure is based on so-called lag-time diffusion methods in which the time to achieve steady-state oxidation is measured using *in situ* ellipsometry. Two different modes of oxidant transport were observed over the range of temperatures investigated (600°-1000°C). At temperatures of 900°C and below, no lag-time was observed, and steady-state oxidation was seen at the outset of oxidation. At 1000°C, a lag-time was measured which yielded a value for the diffusion constant, $D = 2.3 \times 10^{-13}$ cm²/sec for dry O₂, and this value increased to 2.4×10^{-12} cm²/sec for 1000 ppm H₂ in O₂. This study provides clear evidence for different dominant modes of oxidation at higher and lower oxidation temperatures.

The thermal oxidation of silicon is generally believed to proceed by a process that encompasses a steady state of both diffusion of oxidant through an SiO₂ film and reaction of this oxidant with Si at the Si-SiO₂ interface. This notion of the mechanism of oxidation of silicon is primarily due to the success of the linear-parabolic oxidation model (1) in correlating the silicon oxidation data [see for example early reports in Ref. (2 and 3)]. Virtually every worker in this field has reported parabolic-like behavior, i.e., a decrease in the thermal oxidation rate of silicon with time. This plus the several studies that show that oxidant rather than silicon is the primary transported species (4-6) support the diffusion-reaction model. The statistical treatment of copious oxidation data demonstrate reasonable predictability (less than 10% error) of this model (7-9). However, this model is largely phenomenological and therefore contributes only little to the understanding of the details of the oxidation mechanism.

The present study is aimed toward a better understanding of the details of the transport of oxidant across the SiO₂ film. To probe this transport process, the well-known lag-time diffusion technique (10, 11) for thin membranes has been adapted to measure the oxidant transport during oxidation. Basically, the oxidation process is followed using *in situ* ellipsometry. The analytical technique assumes Fickian diffusion of oxidant through the growing oxide film, i.e., transport in response to a concentration gradient, and can therefore be used to test for true diffusion as the dominant mode of oxidant transport.

In the literature, there exists diffusion data for oxygen in fused silica (12-15). These studies report divergent values for both the diffusion constant, D, (several orders of magnitude) and the activation energy for diffusion (a factor of more than three). Diffusion data for oxygen in thin films of SiO₂ has to my knowledge not been reported. Since it is known (16) that the precise method of preparation (temperature, impurities, etc.) of amorphous materials determines many of the physical properties of the material, it is likely that part of the spread in the existing data is due to differing samples. Therefore, in order to understand the oxidation of Si to form SiO₂ films, it is important that the measurement of diffusion be done for thin SiO₂ films.

The present study presents the lag-time method using ellipsometry. Second, the method is used to obtain

data and then an interpretation is presented. The application of this lag-time technique has enabled the measurement of D associated with oxygen transport in SiO₂ at 1000°C and has shown that a lag-time is not observed at 900°C and below. In addition, it was discovered that linear parabolic kinetics obtain at the outset of oxidation for temperatures of 900°C and below. even for films 1000 nm thick. These experimental results lead to considering alternatives to pure Fickian diffusion as the dominant mode of transport of oxygen through SiO₂ films at the lower oxidation temperatures.

Experimental Procedures

Sample preparation.—The experimental technique utilizes SiO₂ films grown by the oxidation of single crystal silicon. The starting silicon wafers were commercially available chem-mechanically polished 2 Ω-cm p-type with (100) orientation. The silicon was cleaned by a previously outlined procedure (7) and the starting SiO₂ films were prepared via thermal oxidation in dry O₂ at 1000°C.

Ellipsometry.—The automated ellipsometer used for the present study, as well as optical constants and procedures, was previously described (7, 17). The ellipsometer can measure oxide growth while the sample is under oxidation conditions. A typical experiment is started by placing a previously grown SiO₂ film of about 10⁴Å on an Si wafer in the oxidation furnace of the automated ellipsometer. The sample is then heated at 1000°C in either N₂ or Ar for about 20 hr. The film thickness is periodically measured to check the quality of the N₂ or Ar and to provide baseline initial oxide thickness data. At some arbitrary time after outgassing, the temperature is adjusted and equilibrated at the desired value and the ambient is switched to oxygen. Before and after this ambient change, SiO₂ thickness is measured continuously. The experiment is usually continued for more than 40 hr after O₂ is turned on to insure that steady-state oxidation is achieved.

In order to insure that any change noticed by ellipsometry is due to a change in SiO₂ thickness and not perhaps a change in optical absorption, a separate experiment was performed. A disk of optical quality fused silica of about 0.64 cm thick and 2 cm in diameter polished on one face and purposely frosted on the other face was placed in the ellipsometer furnace. The polished surface was monitored by ellipsometry and Δ and Ψ data were taken at various temperatures between 600° and 1000°C before and after switching the ambients. No change in the optical constants of the fused silica was observed. Therefore, any change observed

240

for the thick SiO_2 films on Si would be attributed to a real change in the SiO_2 film thickness. As is shown below, this lag-time experiment performed at or below 900°C yielded initial oxidation rates equal to final rates. i.e., no lag-time. Without the experiment above, this observation could possibly be trivially explained by a difference in the optical constants of the oxide in N_2 or O_2 when the ambient is switched. On the other hand, the rapid attainment of steady-state oxidation at the lower experimental temperatures has mechanistic implications and this is covered in the Discussion section.

The temperature range 1000°-600°C was explored; higher temperatures could not be attained in the present equipment.

Lag-Time Method

The lag-time method (10,11) is based on the existence of a time delay for the transport of a gas through a membrane that initially has zero concentration of the gas. The time delay is measured from the time the temperature-equilibrated membrane contacts the gas ($t = 0$) to the time the gas appears in the other side of the membrane as a steady-state flow.

The lag-time analysis due to Daynes (10) and Barrer (11) is herein adapted to the Si-SiO_2-oxygen system. The key equations and boundary conditions are reproduced below with some modification from the published detailed analysis (10,11). For the Si-SiO_2-oxygen gas system, the boundary conditions are shown in Fig. 1 and the general solution for concentration is

$$C = C_1 + (C_2 - C_1)\frac{X}{L} + \frac{2}{\pi}\sum_{m=1}^{\infty}\left(\frac{C_2\cos n\pi - C_1}{n}\right)$$

$$\left(\sin\frac{n\pi X}{L}\right)\exp\left(\frac{-Dn^2\pi^2 t}{L^2}\right) + \frac{4C_0}{\pi}\sum_{m=0}^{\infty}\frac{1}{(2m+1)}$$

$$\sin\frac{(2m+1)\pi X}{L}\exp\left(\frac{-D(2m+1)^2\pi^2 t}{L^2}\right)\quad [1]$$

where C is the concentration of oxidant at any time. At the Si-SiO_2 interface, $X = 0$ and

$$\left(\frac{\partial C}{\partial X}\right)_{X=0} = \frac{C_2 - C_1}{L}$$

$$+ \frac{2}{L}\sum_{n=1}^{\infty}(C_2\cos n\pi - C_1)\exp\left(\frac{-Dn^2\pi^2 t}{L^2}\right)$$

$$+ \frac{4C_0}{L}\sum_{m=0}^{\infty}\exp\left(\frac{-D(2m+1)^2\pi^2 t}{L^2}\right)\quad [2]$$

The rate at which gas, C_g, emerges from a membrane of unit area at $X = 0$ is given as
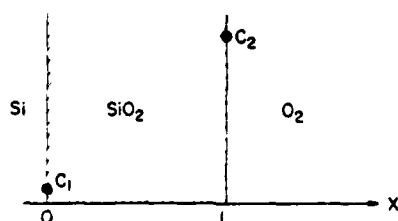
C = C₁ AT X = 0 FOR ALL t

C = C₂ AT X = L

C = C₀ FOR 0 < X < L AT t = 0

Fig. 1. Pictorial representation of Si-SiO₂-O₂ system with boundary conditions.

$$\frac{\partial C_g}{\partial t} = D\left(\frac{\partial C}{\partial X}\right)_{X=0}\quad [3]$$

Now substituting $(\partial C/\partial X)_{X=0}$ from above and solving for the concentration, C_g, by integration the resultant equation is

$$C_g = \frac{D(C_2 - C_1)t}{L}$$

$$+ \frac{2L}{\pi^2}\sum_{n=1}^{\infty}\left(\frac{C_2\cos n\pi - C_1}{n^2}\right)\left(1 - \exp\left(\frac{-Dn^2\pi^2 t}{L^2}\right)\right)$$

$$+ \frac{4C_0 L}{\pi^2}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^2}\left(1 - \exp\left(\frac{-D(2m+1)^2\pi^2 t}{L^2}\right)\right)\quad [4]$$

This concentration will increase toward a steady state as $t \to \infty$ and C_g approaches the line

$$C_g = \frac{D}{L}\left[(C_2 - C_1)t + \frac{2L^2}{D\pi^2}\sum_{n=1}^{\infty}\left(\frac{C_2\cos n\pi - C_1}{n^2}\right)\right.$$

$$\left. + 4\frac{C_0 L^2}{\pi^2 D}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^2}\right]\quad [5]$$

$$= \frac{D}{L}\left[(C_2 - C_1)t - \frac{C_2 L^2}{6D} - \frac{C_1 L^2}{3D} + \frac{C_0 L^2}{2D}\right]\quad [6]$$

The intercept of this line with the time axis yields the lag-time, τ

$$\tau = \frac{1}{(C_2 - C_1)}\left[\frac{C_2 L^2}{6D} + \frac{C_1 L^2}{3D} - \frac{C_0 L^2}{2D}\right]\quad [7]$$

If the actual experiment commences with no oxidant in the film, $C_0 = 0$, and if the reaction at Si-SiO_2 interface is fast relative to tranport then $C_1 \approx 0$, then the lag-time expression reduces to the equation

$$\tau = \frac{L^2}{6D}\quad [8]$$

The condition of $C_0 = 0$ is achieved by a thorough outgas of the sample at 1000°C overnight and it is assumed that the network oxygen does not contribute to oxidation. The condition $C_1 \approx 0$ is consistent with the linear-parabolic model.

The steady-state condition is observed by the conformity of the thickness-time data to the linear parabolic rate law

$$t = AL + BL^2\quad [9]$$

where t is the oxidation time for a film thickness L, and A and B are the reciprocals of the linear and parabolic rate constants, respectively.

This simplified linear-parabolic equation can be used to test for linear-parabolic kinetics in the present study. However, where the values of the rate constants are of interest, it is necessary to utilize the physical model related equation (2,7)

$$t - t_0 = A(L - L_0) + B(L^2 - L_0^2)\quad [10]$$

where L_0 and t_0 define the lower limit of applicability for the linear-parabolic model or an oxide of thickness L_0 at the start of an experiment

Equation [9] is easily linearized and linear least squares analysis of the t and L data according to L vs. t/L yield slope and intercept values that are used to extrapoate to the time axis at the initial oxide thickness value. This intersection yields the τ value used with the average L value in Eq. [8] to obtain D.

To insure that steady state is achieved, oxidation is carried out for more than 40 hr. Thickness-time data from the last 10 hr are analyzed as steady-state values. However, prior to any data analysis, the overall L, t data is sampled in the first and second hours of oxidation and then after 20, 30, and 40 hr. If the lag-time method is applicable, the oxidation rates should be initially low and then increasing toward steady state. This information provides the basis for the lag-time analysis.

Results and Discussion

In order to assess the range of possible lag-time values, τ values were calculated from the divergent D values of Norton (12) and Williams (13) and the results are shown in Table I. It is seen that if Norton's values for D are applicable for SiO_2 films, then τ is virtually unmeasurable by the present technique down to 800°C and marginally measurable down to to 600°C. However, based on Williams' data, τ can be easily measured at 1000°C and below. Therefore, if the measured D values for oxidant in the SiO_2 films used for this study are near Norton's (12) values, we would expect to see no lag-time for the higher temperatures while if Williams' (13) D values obtain, a lag-time would be observed for all temperatures, and for both sets of D values the τ values would be larger for the lower temperatures. The experimental data which consist of thickness-time measurements on films with starting thickness of about 10^4A were first surveyed to determine whether lag-times were observed. As previously stated, this was done by comparing the rates of oxidation at the beginning of the experiment to rates at the end. At least two data sets were obtained at each experimental temperature with substantially the same results. Table II lists the rates from one data set at each temperature for the 1st, 2nd, 20th, 30th, 40th hours of the experiment. It is clearly seen that only at 1000°C is true lag-time behavior observed, i.e., the rate of oxidation increases to a steady-state value. It is also interesting to note that a steady-state rate of about 41 A/min at 1000°C for 10^4A SiO_2 can be calculated from the derivative of Eq. [10] and already published data (18). Given that there could be a maximum of about 10% error in the reported rate constants (7), the rates for the first 2 hr are below the steady-state rate. This calculated result is in agreement with the experimental result to follow and also the steady-state rates of the present study agree with the published values. For the lower temperatures, the initial rates are at least equal to the final rates. There is some data that show a slight trend toward the initial rate being larger than the final rates. However, this type of behavior is anticipated based on linear parabolic kinetics. Representative SiO_2 thickness vs. oxidation time data is shown as Fig. 2a and 3a for 1000° and 800°C, respectively. The solid lines on these plots are drawn to illustrate the change in the slope of the data and are obtained as average slopes calculated from the data for a number of the initial and final data points (see Table II). The 1000°C data have a smaller initial slope than for long times and the opposite is true at 800°C.

Figures 2b and 3b show plots of L vs. t/L at 1000° and 800°C, respectively. The solid lines on these plots are the result of fitting the steady-state data to Eq. [9]. It is clear that the 800°C data are well represented

Table II. Comparison of initial and final rates of oxidation

T (°C)	Oxidation rates (A./hr)				
	1st	2nd	20th	30th	40th
1000	33	57	58	44	44
900	20	18	20	21	21
800	10	9	9	9	8
700	2	2	2	3	2
600	0.7	0.5	0.4	0.3	0.3

by the linear-parabolic model and this was also typical for the 900°, 700°, and 600°C data. The 1000°C data show the expected deviation for shorter times which is indicative of the existence of a lag-time. The 1000°C data were analyzed by linear least squares fitting of L vs. t/L at long times and then finding the intersection of this line with the t/L axis. The intersection is τ/L_0 where L_0 is the starting oxide thickness. τ is then used in Eq. [8] to calculate a value for the diffusion constant. $D = 2.3 \times 10^{-13}$ cm²/sec at 1000°C in dry O_2. This value is about one order of magnitude larger than Williams' (13) value at 1000°C. The value for the lag-time is $\tau = 8490$ sec. It is interesting to compare this value for τ with the value of 8804 sec calculated from the formula derived by considering the approach to steady-state flow through a plane sheet by Crank (19)

$$\frac{D\tau}{L^2} \approx 0.45 \qquad [11]$$

From the relationship between permeability, P, diffusivity, D, and solubility, S

$$S = \frac{PL}{D} \qquad [12]$$

which applies when Henry's law is obeyed, the solubility can be calculated. The permeability is the steady-state oxidant flux which is calculated from the rate of oxidation. At 1000°C the permeability is about $P = 2.7 \times 10^{-12}$ O_2's cm⁻² sec⁻¹ for 1 atm O_2 pressure. This yields a value for solubility of $S = 1.2 \times 10^{22}$ cm⁻³ which is considerably higher than Deal's (2) value of 5.2×10^{16} cm⁻³ calculated using the relationship

$$S = \frac{k_{PAR} \cdot 2.25 \times 10^{22}}{2D} \qquad [13]$$

and Norton's (12) value for D, where k_{PAR} is the parabolic rate constant. The reason for the large difference in S values in the present study and Deal's study (2) is due primarily to differences in D. Using the value for D obtained in this study and Deal's k_{PAR} (2), S is calculated to be 1.6×10^{21} cm⁻³.

To determine the effect of H_2O on D, 1000 ppm was added to dry O_2 by a previously outlined procedure (20). The data at 1000°C showed a lag-time, and a value of $D = 2.4 \times 10^{-12}$ cm²/sec or about a tenfold increase in D with 1000 ppm H_2O was obtained. From previous studies of the affect of H_2O on the rate of Si oxidation (20), we reported less than a twofold increase in rate with 1000 ppm H_2O in O_2. From Eq. [12], if D increases by a factor of 10 and P by a factor of 2 then the solubility of oxidant in the presence of 1000 ppm H_2O must actually decrease by a factor of five.

Table I. Literature values of D and calculated values for the lag-time, τ

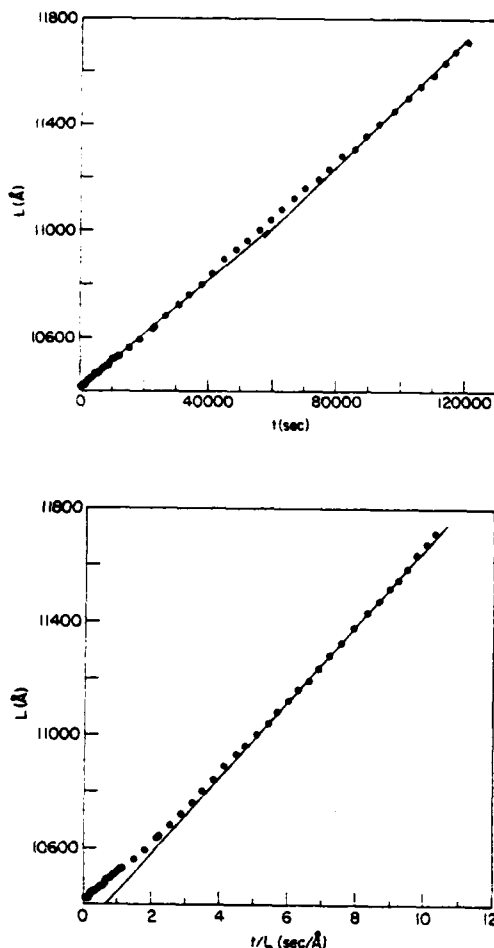|	T (°C)				
	1000	900	800	700	600
D (cm²/sec), from Norton (12)	6×10^{-9}	3×10^{-9}	9×10^{-10}	3×10^{-10}	3×10^{-11}
τ (sec) for 1μ SiO_2	0.3	0.6	2	7	32
D (cm²/sec), from Williams (13)	2×10^{-14}	9×10^{-15}	3×10^{-15}	7×10^{-16}	1×10^{-16}
τ (sec) for 1μ SiO_2	8×10^4	2×10^5	6×10^5	2×10^6	2×10^7

Fig. 2. Thickness, L, and time, t, oxidation data for dry O_2 at 1000°C: (a, top) as L vs. t, (b, bottom) L vs. t/L.

This result may indicate that, although trace amounts of H_2O enhance diffusion of oxidant so that the rate of oxidation increases, the H_2O ties up many active sites in SiO_2 thereby lowering the solubility of O_2.

In terms of the mechanism of oxidation, the present results demonstrate the existence of a change in the dominant mechanism for oxidant transport near 1000°C. Above this temperature, the transport can be considered Fickian with undissociated oxygen as the likely transported species. This latter contention is based on the observation that Henry's law is obeyed at high temperatures (2). The question of whether any transported species are charged is not addressed in this study. Further evidence for the change in predominant mechanism of oxidation was obtained from an earlier study (18) in which considerable curvature in Arrhenius plots involving both linear and parabolic rate constants was reported. In that study (18), the higher temperatures yielded a lower activation energy for k_p in agreement with earlier studies (2) that also emphasized oxidation temperatures above 1000°C. The lower temperatures yielded activation energies for k_p more than 50% larger (18). The mechanism for oxidation below 1000°C is likely to involve a superposition of Fickian diffusion which is predominant at higher temperatures with at least one additional mode of transport. Excluding models which involve charged species, two kinds of mechanisms come to mind which would explain the rapid attainment of steady-state oxidation. One model considers the migration of atomic oxygen from network position to network position.
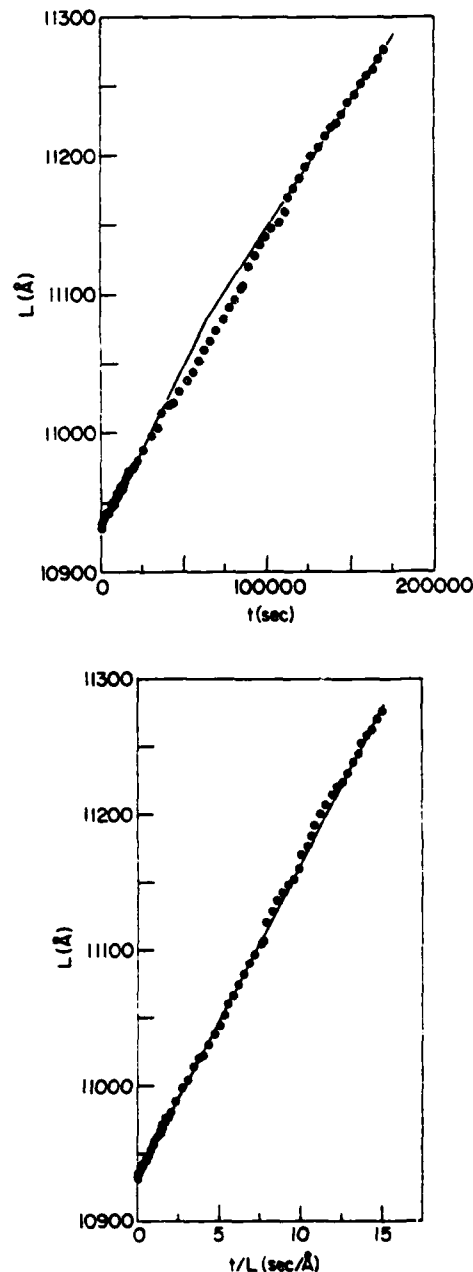

Fig. 3. Thickness, L, and time, t, oxidation data for dry O_2 at 800°C: (a, top) as L vs. t, (b, bottom) L vs. t/L.

Such a mechanism does not require long distance migration of oxygen to initiate oxidation. Rather, when one atom of oxygen is taken into the network at the gas-SiO_2 interface, an atom then becomes available at the Si-SiO_2 interface for oxidation. This process takes place on a time scale of atomic vibrations and would therefore explain the rapid attainment of steady-state behavior. However, the recent tracer studies of Rosencher et al. (21) and Pfeffer and Ohring (22) demonstrate that oxygen is transported across the oxide with little or no interaction with the network. Another possibility is to consider the flow of oxidant in micropores. If micropores exist in sufficient numbers and if the pores penetrate to the Si-SiO_2 interface then very slow diffusion at low temperatures may be short circuited by the flow of oxidant in micropores. Previous studies on thin SiO_2 films yielded some indirect evi-

dence for the existence of micropores in SiO_2 films (23) and recent transmission electron microscopy studies by Gibson and Dong (24) have shown micropores of about 10A to exist in dry O_2 grown SiO_2 films. Further work on a micropore model is necessary and in progress, but it is interesting to note that if Knudsen-Poiseuille flow is considered in micropores (25) in steady state with linear surface kinetics a linear-parabolic type rate law can be derived. The fact that lower oxidation temperatures yield higher density SiO_2 (26, 27) probably also contributes to a decrease in the importance of simple diffusion and perhaps renders a different transport mechanism dominant at lower oxidation temperatures.

Meek (28) has attempted to explain the large discrepancy between Norton (12) and Williams (13) measurements of D for oxygen in fused silica based on the differences in the experimental techniques, *viz.*, Norton measures chemical diffusion based on a lag-time while *Williams measures the transport of a radioisotope*. According to Meek's interpretation, Williams measured value of D is necessarily smaller than Norton's due to the fraction of time the radioisotope spends in exchanging with the network rather than in the interstices of SiO_2. Recently, however, two definitive studies (22, 29) have appeared in the literature which show that the transported oxygen does not exchange to any measurable extent with the SiO_2 network. Therefore, it is now clear that Meek's argument does not apply and the large differences in the D data can be explained either trivially by experimental errors or by real differences in the SiO_2 itself. It is quite clear that both Pfeffer and Ohring (22) and Rigot *et al.* (29) find large differences in the behavior of oxygen in SiO_2 when H_2O is present and it is not unreasonable to speculate that impurities and perhaps even fictive temperature and other preparation conditions lead to the large reported differences in D values.

In summary, a method to measure diffusion of oxidant in growing SiO_2 films has been developed. This method is based on Fickian diffusion through a membrane and therefore can be used to test for diffusion conditions.

Diffusion of O_2 in SiO_2 as the mode of oxidant transport has been found as the dominant mechanism at 1000°C and a value of $D = 2.3 \times 10^{-13}$ cm²/sec has been measured. This value increases by an order of magnitude with 1000 ppm H_2O in the O_2. At lower temperatures, a different mode of oxidant transport is dominant. This mode is characterized by the instantaneous achievement of steady state linear-parabolic oxidation kinetics.

Acknowledgments

REFERENCES

1. U. R. Evans, "The Corrosion and Oxidation of Metals," Chap. XX, Edward Arnold Ltd., London (1960).
2. B. E. Deal and A. S. Grove, *J. Appl. Phys.*, **36**, 3770 (1965).
3. W. A. Pliskin, *IBM J. Res. Dev.*, **10**, 198 (1966).
4. M. M. Atalla, in "Properties of Elemental and Compound Semiconductors," H. Gatos, Editor, p. 163, Interscience Inc., New York (1960).
5. J. R. Ligenza and W. G. Spitzer, *J. Phys. Chem. Solids*, **14**, 131 (1960).
6. W. A. Pliskin and R. P. Gnall, *This Journal*, **111**, 872 (1964).
7. E. A. Irene and Y. J. van der Meulen, *ibid.*, **123**, 1380 (1976).
8. A. G. Revesz and R. J. Evans, *J. Phys. Chem. Solids*, **30**, 551 (1969).
9. M. A. Hopper, R. A. Clarke, and L. Young, *This Journal*, **122**, 1216 (1975).
10. H. A. Daynes, *Proc. R. Soc. London, Ser. A*, **97**, 286 (1920).
11. R. M. Barrer, "Diffusion In and Through Solids," Chap. I, Cambridge University Press, London (1951).
12. F. J. Norton, *Nature (London)*, **171**, 701 (1961).
13. E. L. Williams, *J. Am. Ceram. Soc.*, **48**, 190 (1965).
14. R. Haul and G. Dümgen, *Z. Electrochem.*, **66**, 636 (1962).
15. E. W. Sucov, *J. Am. Ceram. Soc.*, **46**, 14 (1963).
16. R. Roy, *J. Non-Cryst. Solids.* **3**, 33 (1970).
17. Y. J. van der Meulen and N. C. Hien, *J. Opt. Soc. Am.*, **64**, 804 (1974).
18. E. A. Irene and D. Dong, *This Journal*, **125**, 1146 (1978).
19. J. Crank, "Mathematics of Diffusion," p. 49, Oxford University Press, London (1956).
20. E. A. Irene and R. Ghez, *This Journal*, **124**, 1757 (1977).
21. E. Rosencher, A. Straboni, S. Rigo, and G. Amsel, *Appl. Phys. Lett.*, **34**, 254 (1979).
22. R. Pfeffer and M. Ohring, *J. Appl. Phys.*, To be published.
23. E. A. Irene, *This Journal*, **125**, 1708 (1978).
24. J. M. Gibson and D. Dong, *ibid.*, **127**, 2722 (1980).
25. R. M. Barrer, "Diffusion In and Through Solids," Chap. II, Cambridge University Press, London (1951).
26. E. A. Taft, *This Journal*, **125**, 968 (1978).
27. E. A. Irene, D. Dong, and R. J. Zeto, *ibid.*, **127**, 396 (1980).
28. R. L. Meek, *J. Am. Ceram. Soc.*, **56**, 341 (1973).
29. S. Rigo. F. Rochet, A. Straboni, and B. Agius, in "International Conference on Physics of MOS Insulators," North Carolina State University, Raleigh, NC, June 1980.